THE RESPONSE OF SINGLE AUDITORY-NERVE FIBERS IN THE CAT TO SINGLE TONES: SYNCHRONY AND AVERAGE DISCHARGE RATE

by

Don Herrick Johnson

S.B., Massachusetts Institute of Technology
(1970)

S.M., Massachusetts Institute of Technology
(1970)

E.E., Massachusetts Institute of Technology
(1971)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Sept. 1974

Signature of Author .....
Department of Electrical Engineering, August 12, 1974

Certified by .....
Thesis Supervisor

Accepted by .....
Chairman, Departmental Committee on Graduate Students

ARCHIVES 1
MASS. INST. TECH.
DEC 5 1974

THE RESPONSE OF SINGLE AUDITORY-NERVE FIBERS IN THE CAT TO SINGLE TONES: SYNCHRONY AND AVERAGE DISCHARGE RATE

by

Don Herrick Johnson

Submitted to the Department of Electrical Engineering on August 12, 1974 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

ABSTRACT

Responses to single-tone stimuli were recorded from single auditory-nerve fibers in the Dial-anesthesized cat. Responses synchronized to the individual cycles of the stimulus were measured for tone frequencies below 5 kHz and the degree of this synchronization diminished for frequencies above 1 kHz. The reduction of the synchronized response with increasing stimulus frequency is not attributable to limitations in measurement techniques. The average rate of discharge was also measured. Changes in the average discharge rate are not necessarily correlated with changes in synchronization.

The rate of discharge, r(t), of the response to single tones was found to be well approximated by the expression $r(t) = \exp\{a_o + a_1 \cos 2\pi(ft+\theta_1) + a_2 \cos 2\pi(2ft+\theta_2)\}$. As the level of the stimulus was increased, the parameters a_1 and a_2 increased with a_2 proportional to a_1^2 over a part of the response range. An analytical model describing the dependence of the parameters of r(t) upon the parameters of the stimulus was developed. The output of the model closely resembles measured single-tone response patterns.

Some aspects of the responses of auditory-nerve fibers to two tones were also studied. The synchronized response to one tone is suppressed by the simultaneous presentation of a second tone (two-tone synchrony suppression) if a synchronized response occurs to each tone when presented alone. If f_1 and f_2 are the frequencies of the tones, components at the distortion frequencies $f_1 + f_2$ and $f_2 - f_1$ were also measured in the response pattern. At low levels, the amplitudes of these components were found to be equal for $f_1 + f_2$ less than 2 kHz.

In response to two-tone inputs, the output of the model generated for single-tone responses exhibited two-tone synchrony suppression and contained frequency distortion components. The amplitudes of the components synchronized to $f_1 + f_2$ and $f_2 - f_1$ were equal to each other and were also equal to the amplitudes of the corresponding components found in the empirical data.

THESIS SUPERVISOR: Thomas F. Weiss

TITLE: Associate Professor of Electrical Engineering

ACKNOWLEDGMENTS

Through most of the time spent on my doctoral research, financial support was provided by a National Institutes of Health Public Health Service Traineeship (Training Grant No. 5 TOl GM01555-06). Professor William M. Siebert administered the training grant and I am grateful to him for making this traineeship available.

The experiments for this research were performed at the Eaton-Peabody Laboratory of Auditory Physiology. The efficiency and capability of the members of this laboratory contributed to an atmosphere conducive to scientific investigation. The surgical preparation of the animals was performed by Miss Elizabeth Marr, Miss Susan Mrose, and Miss Catherine Pike. The figures for this thesis were drawn by Elizabeth Marr; the clarity and neatness of the figures reflect her drafting skills. This thesis was typed by Mrs. Eve Golden; the reader can attest to her skills.

Many of the routine computations were performed on the computer facility run by Professor Louis D. Braida. He allowed ready access to the computer for which I am grateful.

While virtually all members of the Eaton-Peabody Laboratory provided food for thought as well as helpful criticism, a few deserve special attention. Discussions with Doctor Julius L. Goldstein during the formative stages of this research were very valuable. He gave freely of his thoughts and opinions; they have influenced my conception of how to approach the construction of a model. Professor H. Steven Colburn served as a thesis reader and also as my graduate counselor.

He was always available for a discussion of technical problems and his advice was well given. The Eaton-Peabody Laboratory is directed by Doctor Nelson Y.-S. Kiang; he formally served as a thesis reader. His contributions to this research were innumerable. He was always willing to discuss matters and provided guidance in scientific and physiological matters. The basic form of the model developed here was originally suggested by him. Professor Thomas F. Weiss served as the research supervisor. His influence upon this research goes well beyond his formal responsibilities. His advice concerning all phases of this research was well given, and is deeply appreciated by me. Finally, there is my wife, Carol. She has continually demonstrated patience and understanding during the years and has provided the environment necessary for the completion of this research. All of this was accomplished while caring for our two children. Her role cannot be overestimated.

TABLE OF CONTENTS

Page

Page

Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The assimilation of acoustic information by mammals begins with the mechanical processes that take place in the peripheral auditory system. Sound impinging upon the external ear sets the middle-ear ossicles in motion, providing the mechanical input to the cochlea where mechanical-to-neural transduction occurs. Individual auditory-nerve fibers exhibit probabilistic sequences of spikes whose rate reflects the spectral-temporal properties of the sound stimulus. The discharge patterns of these nerve fibers provide the only information concerning the acoustic environment to the central nervous system of the animal.

Investigators have measured the responses of single auditory-nerve fibers to a variety of stimuli. In particular, the response pattern to tonal stimuli have provided much insight into the response characteristics of the peripheral auditory system. Each auditory-nerve fiber is most sensitive to tones in a restricted range of frequencies. As the stimulus level is increased, the average rate of discharge increases, and for low-frequency tones, the discharges become increasingly synchronized with the individual cycles of the stimulus. However, the response to a single tone has not, heretofore, been quantitatively characterized by an analytical model. This study is concerned with precise measurements of the rate of discharge of single auditory-nerve fibers in the anesthesized cat responding to single tone. Through an analysis of the average-rate response

and the synchronized response, a model is then developed to characterize the single-tone response pattern.

CHAPTER II

ANATOMY AND PHYSIOLOGY OF THE PERIPHERAL AUDITORY SYSTEM

A schematic diagram of the gross anatomical features of the human peripheral auditory system is presented in Figure 2.1. The peripheral auditory system of most mammals is similar to that of the human. A functional block diagram of the auditory periphery is found in Figure 2.2.

2.1 The External Ear

The structures comprising the external ear are the head, the pinna (auricle), and the external auditory meatus (ear canal) terminated by the tympanic membrane (eardrum). Measurements have been made which compare sound pressure level at the tympanic membrane $p_d(t)$ with sound pressure level at the tragus $p_t(t)$ (Wiener and Ross, 1946; Wiener et al., 1966). According to these measurements, the acoustic transmission properties of the external auditory meatus vary with frequency, having a broad, dominant peak in the 4-5 kHz range.

Figure 2.1. Schematic Drawing of the Peripheral Auditory System

The gross structures comprising the human peripheral auditory system are depicted in this schematic drawing (after von Békésy and Rosenblith, 1951). The cochlea is shown uncoiled and the cochlear partition is represented by only a line. Variables superimposed upon the drawing are:

p_t – sound pressure at the tragus or, equivalently, the entrance to the external auditory meatus

p_d – sound pressure at the tympanic membrane

x_s – linear displacement of the stapes

$y(x)$ – displacement of the cochlear partition as a function of x, longitudinal distance along the cochlear partition.

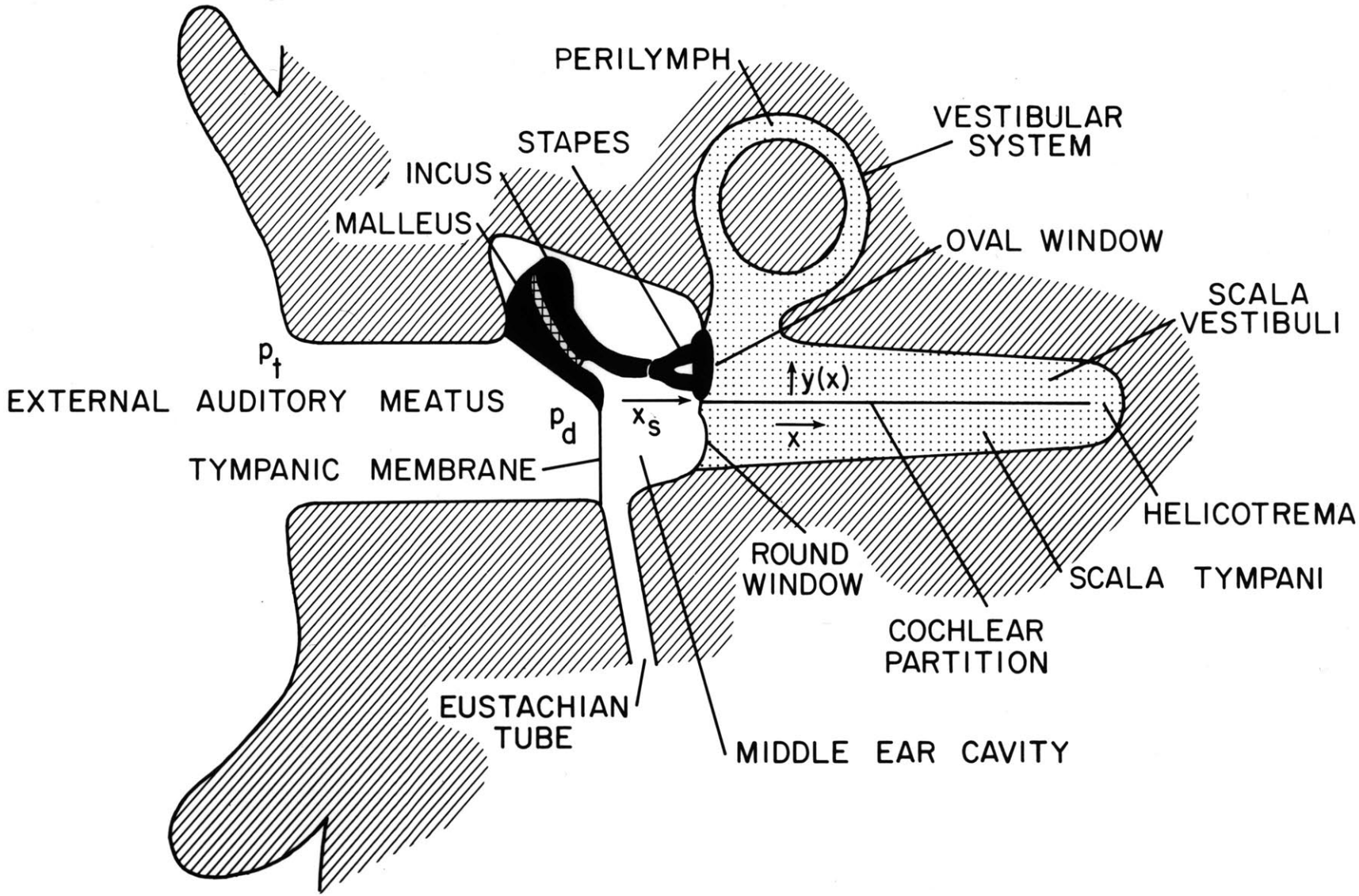

PERILYMPH
VESTIBULAR SYSTEM
STAPES
INCUS
MALLEUS
OVAL WINDOW
SCALA VESTIBULI
p_t
EXTERNAL AUDITORY MEATUS
p_d
x_s
y(x)
x
TYMPANIC MEMBRANE
HELICOTREMA
ROUND WINDOW
SCALA TYMPANI
COCHLEAR PARTITION
EUSTACHIAN TUBE
MIDDLE EAR CAVITY

Figure 2.2. Block Diagram of the Functional Behavior of the Peripheral Auditory System

The conceptual steps in the transformation from a sound stimulus to auditory-nerve fiber's discharge pattern are indicated in this block diagram. Each block is intended to represent a transformation from its input to its output variable. The labelled variables are:

$p_t(t)$ – sound pressure at the tragus

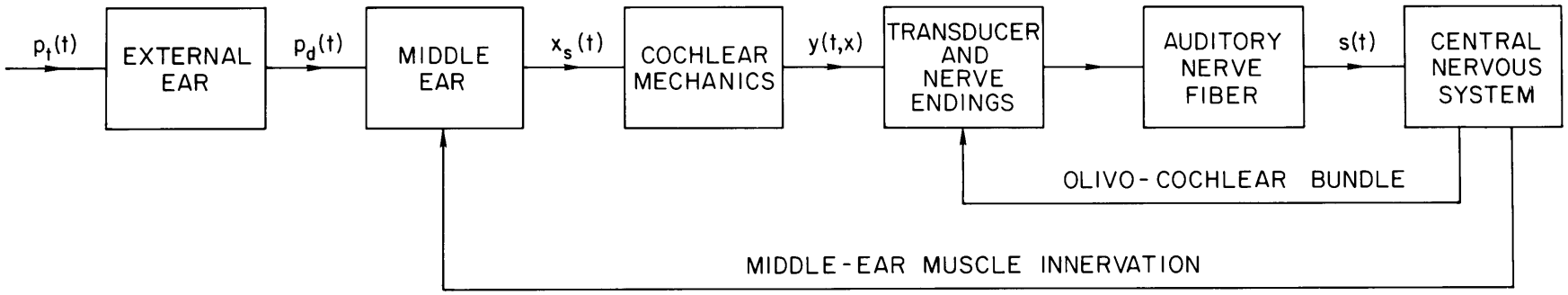
$p_d(t)$ – sound pressure at the tympanic membrane

$x_s(t)$ – linear displacement of the stapes

$y(t, x)$ – displacement of the cochlear partition as a function of longitudinal distance x.

$s(t)$ – spike pattern of an auditory-nerve fiber.

The innervation of middle-ear muscles and the innervation of the hair cells and afferent nerve endings by the olivo-cochlear bundle are indicated.


$p_t(t)$
EXTERNAL EAR
$p_d(t)$
MIDDLE EAR
$x_s(t)$
COCHLEAR MECHANICS
$y(t,x)$
TRANSDUCER AND NERVE ENDINGS
AUDITORY NERVE FIBER
$s(t)$
CENTRAL NERVOUS SYSTEM
OLIVO-COCHLEAR BUNDLE
MIDDLE-EAR MUSCLE INNERVATION

2.2 The Middle Ear

The middle ear converts sound pressure at the tympanic membrane, $p_d(t)$, into mechanical displacements of the stapes, $x_s(t)$, in the oval window of the cochlea. The entire tympanic membrane vibrates in-phase for low-frequency tones but its motion becomes more complex for frequencies higher than a few kHz (Békésy, 1960: 101-102; Tonndorf and Khanna, 1960). Motion of the tympanic membrane sets in motion the three middle ossicles (bones): the malleus (hammer), the incus (anvil) and the stapes (stirrup).

Guinan and Peake (1967) measured the motion of the middle-ear ossicles in anesthesized cats. They described the motion of the stapes as "piston-like": the footplate of the stapes moved in and out of the oval window. When a sinusoidal sound pressure $p_d(t)$ was applied, the motion of the middle-ear ossicles was found to be sinusoidal for sound pressure levels less than 140 dB SPL. These results suggest that the motion of the stapes might be linearly related to the sound pressure at the tympanic membrane. If so, a transfer function relating $x_s(t)$ to $p_d(t)$ can be meaningfully defined. When the bulla is opened and the septum removed, this transfer function is lowpass in nature with a cutoff frequency of approximately 1.5 kHz.

Two middle-ear muscles, the tensor tympani and the stapedius, are attached to the middle-ear ossicles. The tensor tympani muscle is attached to the malleus while the stapedius muscle is attached to the stapes. Bilateral contractions of these muscles can occur with more generalized muscle movements of the animal or intense sound stimulation to either

ear (Møller, 1962; Carmel and Starr, 1963; Borg, 1972). Apparently the middle-ear muscles are relaxed when the animal is under barbituate anesthesia (Carmel and Starr, 1963); furthermore, these muscles do not seem to affect the transmission properties of the middle ear when they are relaxed (Simmons, 1959; Dallos, 1970).

2.3 The Cochlea

The sequence of events in the cochlea leading from motion of the stapes to discharges in auditory-nerve fibers is not well understood. Accessibility to cochlear structures is limited; measurements of cochlear activity are, for the most part, only beginning to provide necessary data.

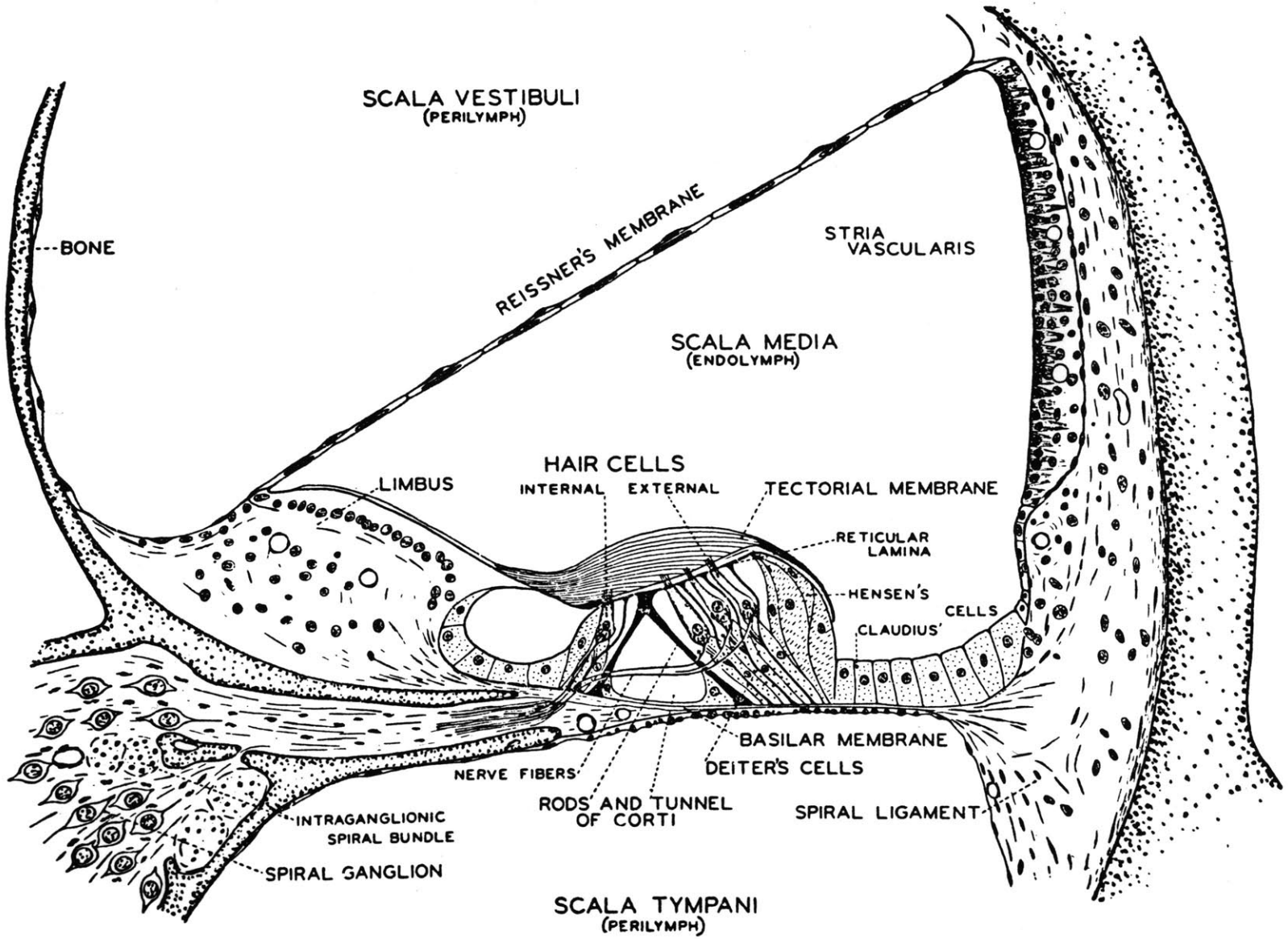
2.3.1 Anatomy of the cochlea and the auditory nerve

The mammalian cochlea contains three fluid-filled cavities that are coiled about a central bony core (the modiolus). The cochlea is shown diagramatically in Figure 2.1 as if it were uncoiled. The cochlea can be described grossly as a tube within a tube. The central tube, the cochlear partition, separates the main tube into two compartments: scala vestibuli and scala tympani. The dimensions of the cochlear scalae and the cochlear partition vary along the length of the cochlear spiral. The cochlear partition ends near the apex, allowing scala tympani and scala vestibuli to communicate through an opening, the helicotrema.

Figure 2.3 depicts a cross-section through one portion of the cochlear spiral. The cochlear partition is bounded by Reissner's membrane and the basilar membrane, and contains the sensory organ, the organ of Corti. The fluid-filled space of the cochlear partition is called scala media; scala media is filled with endolymph. Scala tympani and scala vestibuli are filled with a fluid called perilymph. Endolymph and perilymph differ in their chemical composition (Smith et al., 1954; Citron et al., 1956).

Figure 2.3. Drawing of a Cross-section of the Cochlear Partition

The cochlear scalae and the structures comprising the cochlear partition are indicated in this drawing (Davis, 1959). This drawing was made from a cross-section of the cochlea of a guinea pig.


SCALA VESTIBULI
(PERILYMPH)
BONE
REISSNER'S MEMBRANE
STRIA VASCULARIS
SCALA MEDIA
(ENDOLYMPH)
HAIR CELLS
INTERNAL
EXTERNAL
LIMBUS
TECTORIAL MEMBRANE
RETICULAR LAMINA
HENSEN'S CELLS
CLAUDIUS'
BASILAR MEMBRANE
DEITER'S CELLS
NERVE FIBERS
RODS AND TUNNEL OF CORTI
SPIRAL LIGAMENT
INTRAGANGLIONIC SPIRAL BUNDLE
SPIRAL GANGLION
SCALA TYMPANI
(PERILYMPH)

The organ of Corti contains the sensory cells: the inner and outer hair cells. The inner hair cells are separated from the outer hair cells by the tunnel of Corti. Along the length of the cochlear partition, there is a single row of inner hair cells and three rows of outer hair cells (cat). Cilia (hairs) protrude from the tops of the hair cells and at least some cilia are imbedded in the tectorial membrane (Kimura, 1966). Synaptic contacts with the afferent fibers are found at the base of the hair cells.

The afferent innervation of the sensory hair cells is provided by a division of the VIIIth cranial nerve, the auditory nerve. The innervation pattern of the auditory nerve in the cochlea is not known completely. Two types of afferent fibers are found: radial fibers and spiral fibers (Lorente de Nó, 1937). Radial fibers, comprising 90-95% of the total (Spoendlin, 1969), innervate the inner hair cells. One radial fiber innervates only one inner hair cell, but one inner hair cell makes contact with about 20 radial fibers (Spoendlin, 1972). The spiral fibers, on the other hand, enter the cochlea, turn basalwards and innervate many outer hair cells. It may be possible that spiral fibers may also make contact with the inner hair cells (Perkins, 1973).

The cell bodies of the auditory-nerve neurons form the spiral ganglion, which lies in a spiral canal (Rosenthal's canal) in the temporal bone. These neurons are bipolar; their central processes comprise the auditory nerve which projects to the ipsilateral cochlear nucleus. In the cat, the auditory-nerve contains approximately 50,000 fibers having diameters ranging from 2-6 μ (Gacek and Rasmussen, 1961).

An efferent innervation is also present in the cochlea. The bundle of efferent fibers is termed the olivocochlear bundle (OCB). The OCB

originates from cells near the superior olivary complexes of the ipsilateral and contralateral sides. The number of nerve fibers in the OCB at the point it joins the auditory nerve is small, perhaps 500 (Rasmussen, 1960). Extensive ramifications of the efferent supply occur in the cochlea (Spoendlin, 1966: 38-41). Efferent fibers terminate predominantly on the bases of outer hair cells and on the unmyelinated terminals of the auditory-nerve fibers under the inner hair cells (Spoendlin and Gacek, 1963).

2.3.2 Cochlear mechanics

The motion of the stapes in the oval window serves to create a sound-pressure difference across the cochlear partition (Nedzelnitsky, 1974). This pressure difference results in motion of the cochlear partition. The displacement of the cochlear partition in the apical portion of the cochlea was measured by von Bĕkĕsy (1960: 446-469, 500-510) in several animals, including human and guinea pig. More recently, motion of the basilar membrane in the more basal turns has been measured in the guinea pig (Johnstone et al., 1970; Kohllöffel, 1972; Wilson, 1974) and in the squirrel monkey (Rhode, 1971). These data indicate that the displacement of one point along the basilar membrane $y(t, x)$ exhibits a band-pass frequency-response characteristic to sinusoidal motion of the stapes. This frequency-response characteristic is termed a mechanical tuning curve. The frequency corresponding to the peak of the mechanical tuning curve depends upon x, the distance from the point of measurement to the stapes. Most workers have reported that the motion of the basilar membrane is linearly related to the motion of the stapes. However, Rhode (1971) found evidence of a nonlinear relationship. The exact nature of

this nonlinearity is not certain; however, it seems to be localized to frequencies corresponding to the region of the peak of the mechanical tuning curve.

No measurements have been made of the motion of the smaller structures in the cochlear partition, such as the relative motion of the hair cells and the tectorial membrane. Various authors have suggested that bending the stereocilia of the hair cells provides the necessary stimulus for the initiation of spikes in auditory-nerve fibers (Davis, 1950; Békésy, 1960: 703-710; Dallos, 1973: 196-203).

2.3.2 Cochlear electrophysiology

If a gross electrode is placed on the bone near the round window, a potential can be recorded upon the presentation of an acoustic stimulus. In response to a brief acoustic pulse (click), two components separated in time have been identified in this potential. The early component, termed the cochlear microphonic, has been presumed to be related to the sensory hair cells. The later components, including the N_1-N_2 complex, reflect activity of the auditory-nerve fibers and its projection areas.

There is an extensive literature concerning intracochlear and extracochlear potentials. As this subject is not of primary concern in this thesis, the topic in not pursued further.

2.4 Responses of Auditory-Nerve Fibers to Acoustic Stimuli

Auditory-nerve fibers conduct all-or-nothing neural action potentials (spikes) from the cochlea to the cochlear nucleus. These spikes have nearly identical waveforms; information concerning the acoustic stimulus is presumably encoded in the timing of the spikes rather than in their detailed waveforms. Consequently, these discharge patterns may be modelled as stochastic point processes (Cox and Lewis, 1968). The responses recorded from single auditory-nerve fibers of the cat have been studied systematically by Kiang (1965, 1968). A description of the responses to tonal stimuli is most relevant here.

2.4.1 Spontaneous activity

In the absence of controlled acoustic stimuli, a random, spontaneous sequence of spikes is recorded from every auditory-nerve fiber. The average rate of discharge of this spontaneous activity ranges from less than 1 spike/sec to greater than 100 spikes/sec. The interval histogram provides an estimate of the probability density function of the interval between successive spikes. The interval histogram computed from the spontaneous activity of an auditory-nerve fiber is characterized by a single mode of less than 10 msec and an exponential decay for intervals longer than the mode. Presumably, the lack of short intervals is due to the refractory properties of spike initiation (Moxon, 1967). Statistical tests (Rodieck et al., 1962) indicate that the spontaneous activity of an auditory-nerve fiber can be represented as a renewal process (Cox, 1962).

Furthermore, the spontaneous activities of pairs of auditory-nerve fibers can be described as independent renewal processes (Johnson, 1970).

2.4.2 Single-tone responses

When a single tone is presented as an acoustic stimulus, a particular auditory-nerve fiber will respond only if the stimulus amplitude and frequency lie within the response area of the fiber. As a function of frequency (f), the stimulus level evoking a just-detectable increase in the average discharge rate above the spontaneous rate is termed a tuning curve. The tuning curve of an auditory-nerve unit exhibits a minimum occurring at a frequency characteristic for each unit; this frequency is termed the characteristic frequency or CF. The tuning curve rises sharply for frequencies above and below the CF. The response area of a fiber is the portion of the level-frequency plane lying within the tuning curve. An auditory-nerve fiber's tuning curve has greater frequency selectivity than measured mechanical tuning curves (Evans, 1972; Geisler et al., 1974). The spontaneous discharge rate is not related to the CF of the fiber; consequently, at least these two parameters are needed to specify the fiber physiologically (Kiang et al., 1965: 94-95).

As the level of a tone is increased above its average-rate threshold, the average discharge rate increases over a stimulus range of 20-40 dB to maximum average discharge rate of less than 200 spikes/sec. Generally speaking, the average discharge rate does not change for further increases in level if the sound pressure level at the tympanic membrane is less than 90 dB SPL.

For levels greater than 90 dB SPL, a sharp decrease or dip in the average rate occurs at some stimulus frequencies (Kiang et al., 1969). The average discharge rate returns to a large value within 10 dB above the level at which the dip begins.

If a tone is presented for an extended period of time (minutes), the average rate of discharge of an auditory-nerve unit decreases (adapts) during the presentation interval. After a period of approximately five minutes, the average discharge rate remains relatively constant at a rate greater than the spontaneous rate (Kiang et al., 1965: 73, 78-79).

If the frequency of the tone is less than about 5 kHz, the spikes tend to occur during a preferred portion of the stimulus cycle (Kiang et al., 1965: 79, 81; Rose et al., 1967). For tone frequencies greater than 5 kHz, the synchronized response is small. The exact manner in which the discharges are synchronized with stimulus does not seem to depend upon the refractory properties of the response pattern (Gray, 1966). The mechanism responsible for the lack of stimulus-synchronized discharges at high frequencies is not precisely known. Rose (1967) does not believe it to lie in the refractory properties of the nerve fiber. Anderson (1973) has proposed that jitter in the propagation time of a spike in a fiber accounts for this loss of synchrony. The precision with which a synchronized response could be measured at high frequencies has not been assessed.

Synchronized responses can be detected for levels that are lower than those required to detect a change in the average rate of discharge

(Rose et al., 1971; Littlefield et al., 1972). For larger levels, there is more synchronization of the discharges to the stimulus (Hind et al., 1967). In some instances, discharges tend to occur during two portions of a stimulus cycle rather than one (Gray, 1966). These responses appear to be synchronized to the second harmonic of the stimulus frequency (Kiang, 1974). This response pattern is not due to second-harmonic distortion in the acoustic stimulus (Kiang and Moxon, 1972). This response pattern is termed "peak-splitting" (Kiang, 1974).

The phase angle of the synchronized response has been studied (Pfeiffer and Molnar, 1970; Anderson et al., 1971). The phase of the synchronized response of an auditory-nerve fiber is (approximately) linearly related to tone frequency, the slope of the relationship depending upon the CF of the fiber. In addition, the phase of the response to a fixed-frequency tone may change as a function of stimulus level. Anderson et al., (1971) report phase changes with level of no more than one-quarter of a period for sound pressure levels of the tone at the tympanic membrane of less than 90 dB SPL. For frequencies less than the CF of the fiber, the phase measurements indicate increasing lag with increasing level; above CF, the phase measurements show more lead (Anderson et al., 1971). A similar effect has been noted in measurements of basilar membrane motion (Rhode and Robles, 1974). Kiang et al. (1969) reported that phase changes (about one-half period) can occur for levels near the region of the dip in the average rate.

2.5 Models of Auditory-Nerve Activity

Many models of the stimulus-related activity of auditory-nerve fibers have been developed. For purposes of this thesis, models of tonal response patterns are reviewed here.

The structures of previous models describing single-tone response patterns are similar; the basic structure proposed by Weiss (1966) has prevailed in models of the peripheral auditory system. His model consisted of a cascade of three fundamental elements: (1) a system having a tuned input/output relationship (e.g., a bandpass filter), (2) a transducer which rectifies its input waveform, and (3) a spike generator. Weiss' original model (1966) consisted of a linear, time-invariant, bandpass filter followed by a memoryless* nonlinearity with a threshold-crossing device as the spike generator. The gross features of the single-tone response patterns (an average rate increase and the presence of a synchronized response) are described by this model; however, the detailed characteristics of the synchronization of the discharges to the stimulus tone are not well described (Geisler, 1968). With one exception (Pfeiffer and Kim, 1973), further modeling work has been concentrated upon the last two stages of the model: the transducer and the spike generator.

---

*"Memoryless" is a term implying that the amplitude of a device's output is related to the amplitude of its input without regard to the time-behavior of the input signal.

Siebert (1970) proposed a model having the rate of discharge of the output spike train an exponential function of the stimulus waveform. This model was intended to represent the main features of the single-tone response pattern for low-frequency stimuli. In more detail, letting r(t) be the rate of discharge and $x(t) = X \cos 2\pi ft$ be the output of the tuned filter, then Siebert (1970) and Colburn (1973) proposed:

$$r(t) = r_{sp} \exp\left\{ \frac{X \cdot |G(f)|}{\sqrt{K + X^2}} \cos 2\pi ft \right\} \qquad (2.1)$$

where: G(f) is a linear, first-order, lowpass filter.

r_{sp} is the spontaneous discharge rate.

K is a constant.

This model provides an approximation to the rate of discharge of synchronized responses to a tone and the relationship of average discharge rate to stimulus level (Colburn, 1973). The filter G(f) is intended to represent the decline in the synchronized response with increasing stimulus frequency. However, this model does not have any response unless there is a synchronized response. Consequently, this model mimics auditory-nerve fiber response patterns only at stimulus frequencies less than a few kilohertz.

Schroeder and Hall (1974) focused upon the spike generation process. Their model is capable of mimicking the variation of the average rate with level and some general features of the synchronization of discharges to the stimulus.

Though not formulated explicitly, Rose (1970, 1974) assumes that the relation between discharge rate and the stimulus is described by a half-wave rectifier with a threshold. Using the previously defined notation:

$$r(t) = \begin{cases} x(t) - A & x(t) \geq A \\ 0 & x(t) < A \end{cases} \qquad (2.2)$$

The accuracy of the descriptions of this model has not been quantitatively evaluated.

Pfeiffer and Kim (1973) have proposed a model which incorporates nonlinear behavior into the tuned filter (Kim et al., 1973a). The response of this model to a single tone shows phase shifts with level and saturation of the response at high levels. A detailed comparison of data obtained from auditory-nerve fibers with the predictions of their model has not been reported.

2.6 Outline of Thesis Research

As described earlier, measurements have been made of the synchronized response (Rose et al., 1967) and of the average-rate response (Kiang et al., 1965: 79-83; Kiang et al., 1969; Geisler et al., 1974; Sachs et al., 1974) to a single-tone stimulus. In order to characterize the response to single tones simultaneous measurements of the synchronization and the average rate need to be obtained systematically as functions of stimulus level and frequency.

For the purpose of this thesis, the responses of single auditory-nerve fibers to single tones in the anesthesized cat are obtained for stimulus frequencies less than 5 kHz and stimulus levels less than 80 dB SPL. In this manner, both the synchronized response and the average-rate response are studied. The response measures S_f, the synchronization index, and $\bar{r}$, the average discharge rate are defined and their relation to stimulus parameters studied. These data provide the empirical basis for this study.

The waveform of the rate of discharge r(t) is then approximated by:

$$r(t) = \exp\{a_o + a_1 \cos 2\pi(ft+\theta_1) + a_2 \cos 2\pi(2ft+\theta_2)\} \qquad (2.3)$$

This description depends upon five parameters: a_o, a_1, θ_1, a_2, and θ_2. The values of these parameters are assigned by a maximum likelihood estimation procedure which uses the response measures as its database. The description of equation (2.3) compares favorably with the data.

Based upon the relation of the parameters to the level and frequency of the tone, an analytical model is developed. This model of the function of the peripheral auditory system describes many features of the single-tone response pattern. Some attempt is made to examine the prediction of the model for a two-tone stimulus. Some, but not all, of the two-tone response characteristics measured in auditory-nerve fiber spike data are found in the output of the model.

Chapter III details the surgical and experimental methods, and, in addition, an analysis of the procedure used in obtaining the measurements. In Chapter IV, measurements of the single-tone response pattern are presented and the model developed. In order to extend the model to other stimuli, the synchronized responses to two tones are measured and are compared with predictions of the model in Chapter V. The implications of the data and the model are discussed in Chapter VI.

CHAPTER III

METHODS

3.1 Preparation of the Animal

Adult cats weighing between 1.0 and 5.3 kilograms were used as experimental animals. The cats were anesthesized with intraperitoneal injections of Dial (75 mg/kg of body weight). Booster injections were given if the animal demonstrated a withdrawal reflex to a toe pinch.

The surgical details are found in Kiang et al. (1965: 3-4). The bulla and septum of each ear were opened widely. The middle ear muscles were left intact. The external auditory meati were cut and earbars inserted for the placement of the cat's head in a headholder. The posterior fossa of the skull was opened and the dura was removed to expose the cerebellum.

The preparation was placed in a double-walled chamber constructed for acoustic and vibration isolation and electrical shielding.

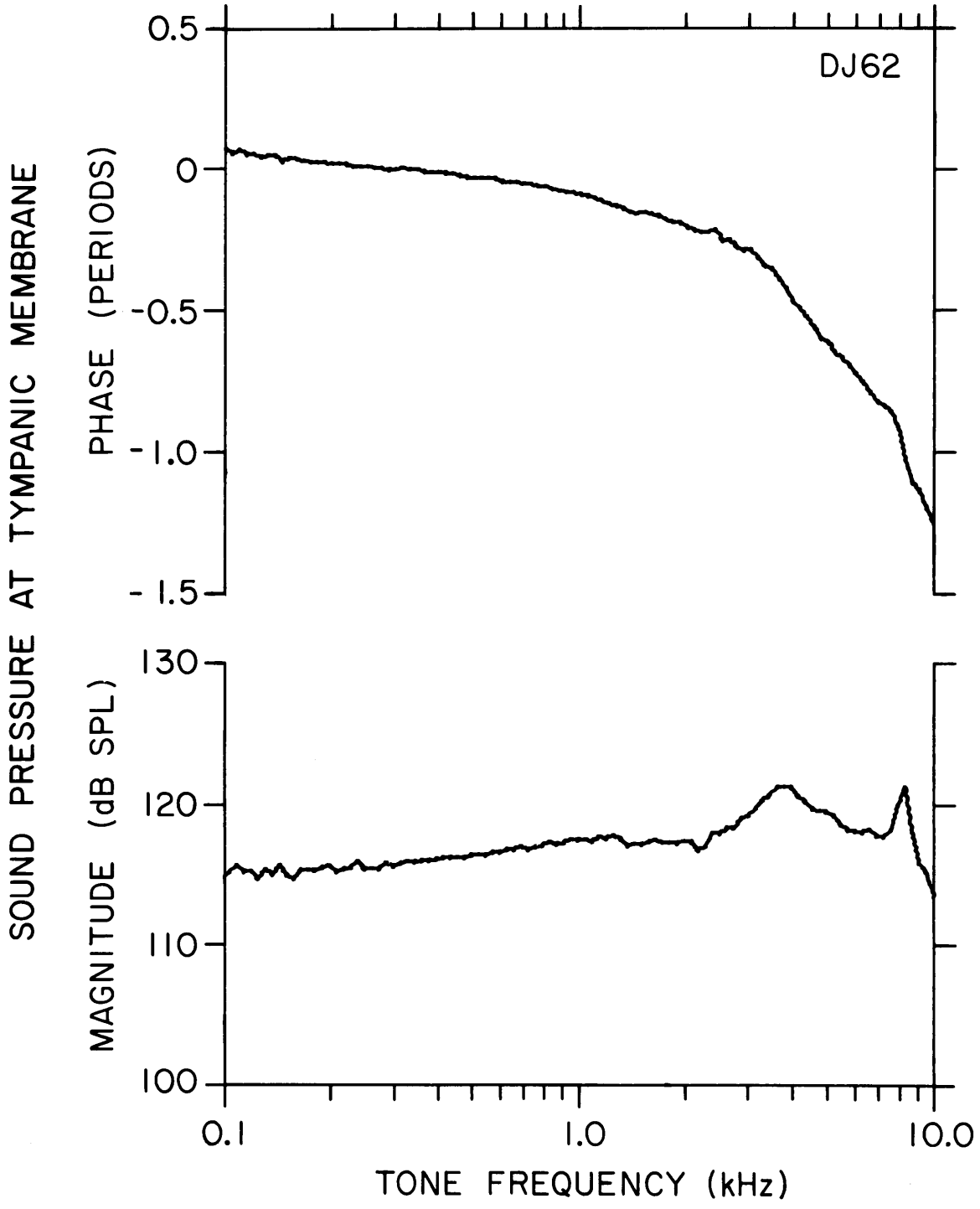
3.2 Stimulus Generation and the Acoustic System

The earphone-probe-microphone assembly (see Figure 2.2, Kiang, et al., 1965) was acoustically sealed into the external auditory meatus near the tympanic membrane. Acoustic stimuli were delivered by applying electrical signals to a one-inch diameter condenser earphone (Brüel and Kjær 4132). The sound pressure near the tympanic membrane was monitored by a probe-microphone. A one-quarter inch diameter condenser microphone (Brüel and Kjaer 4136) served as the monitoring microphone. The magnitude of the transfer function of the probe-microphone was measured prior to each experiment. The magnitude of the transfer function relating the electrical output of the stimulus oscillator to the sound pressure at the tympanic membrane was obtained for each experiment. The level of all tonal stimuli is expressed in dB SPL (rms sound pressure level re 0.0002 dynes/cm^2) at the tympanic membrane. In most experiments, the phase angle of the sound pressure relative to the electric stimulus was not computed. Consequently, the phase of auditory-nerve fiber responses is referred to the electrical output of the stimulus oscillators. Phase is specified in terms of fractions of a period (e.g., a phase of 0.5 is equivalent to 180°) with positive phase corresponding to lead. Phase measurements of the sound pressure at the tympanic membrane show little phase shift (±0.1 period) below 1 kHz (see Figure 3.1).

Tonal stimuli were generated by low-distortion (less than 0.4% total distortion) oscillators. Tone bursts were obtained by passing

Figure 3.1. Magnitude and Phase Calibrations Curves of the Acoustic Stimulus System

The magnitude and phase of the sound pressure level at the tympanic membrane for a constant oscillator voltage are shown as a function of frequency. The level of the sound pressure is expressed in dB re 0.0002 dyne/cm^2 rms (dB SPL). The levels shown here correspond to the maximum sound pressures that can be generated by the acoustic stimulus system. Positive values of the phase denote lead of the sound pressure with respect to the oscillator voltage. These curves consist of data points measured at a density of 50 points for decade of frequency and connected with straight lines.


SOUND PRESSURE AT TYMPANIC MEMBRANE
PHASE (PERIODS)
0.5
0
-0.5
-1.0
-1.5
DJ62
MAGNITUDE (dB SPL)
130
120
110
100
0.1
1.0
10.0
TONE FREQUENCY (kHz)

a sinusoid through an electronic switch gated at a rate of 10/sec. The rise/fall time of the tone bursts was 2.5 msec as measured from baseline to full burst amplitude on an oscilloscope. The tone burst duration was 50 msec as measured between the half-amplitude points of the tone burst. Acoustic clicks were derived from rectangular electrical pulses (100 μsec duration) applied to the earphone at a repetition rate of 10/sec.

3.3 Recording of Responses

3.3.1 Recording of gross cochlear potentials

A wire electrode for measuring gross cochlear potentials was placed on the periostium near the round window. The reference electrode was attached to the headholder. Responses to acoustic clicks as recorded by the wire electrode (gross cochlear microphonic and N_1) were amplified and used to judge the overall sensitivity of the cochlea and the stability of the preparation. The level of a rarefaction-pressure click which evoked a just-noticable N_1 response (as judged on the face of an oscilloscope) is referred to as the visual detection level or VDL. The VDL was monitored throughout an experiment and, for the experiments reported here, the VDL remained constant within 5 dB.

3.3.2 Recording of single-unit responses

Micropipettes filled with 2M KCl were used for the recording of single-fiber responses. Tip resistances, as measured in a 2M KCl bath, ranged between 15 and 50 megohms. The micropipette was placed in the auditory nerve under visual inspection with the aid of an otoscope. A thin, platinum wire was inserted in the shank of the micropipette. A lead attached to the platinum wire led to a cathode follower; the reference electrode was attached to the headholder. The output of the cathode follower was amplified and led outside the chamber for processing and recording. Spikes recorded by the micropipette-amplifier

system were passed through an RC highpass filter (1.5 kHz 3 dB point) to reduce low-frequency noise and to sharpen the spike waveform to make triggering more precise.

3.3.3 Processing of spike trains

The train of spikes recorded by the micropipette, the gross cochlear potential, and the stimulus waveforms were recorded on magnetic tape using a FM tape recorder (Ampex FR-1300) for off-line processing. The spike train was converted to a train of electric pulses via a pulse generator triggered on the initial phase of each spike. PST (post-stimulus time) histograms were computed by a LINC computer and stored on LINCtape during the experiment. PST histograms synchronized to the period of a sinusoidal stimulus (i.e., a period histogram) were referenced to the positive-going zero-crossing of the output of the stimulus oscillator. Interval histograms were computed off-line by the LINC computer from the data recorded during the experiments. These histograms were also stored on LINCtape.

3.4 General Protocol of the Experiments

The search stimulus consisted of clicks presented at a level less than -40 dB re 100 volts peak into the earphone. Once a unit was isolated, the spontaneous discharge rate was estimated by counting the number of spikes occurring in a time interval (15 or 30 seconds in duration). A tuning curve was then measured automatically under computer control (Kiang et al., 1970). To obtain each point in the tuning curve, the frequency and level of the tone bursts were adjusted to evoke, on the average, one more spike during the tone burst than occurred in the silent interval. The protocol for the sequence of stimulus presentation after the tuning curve measurement varied from one experiment to another. The protocol for two-tone experiments is presented in section 5.2.

In experiments on responses to single-tone stimuli, a tone of constant stimulus parameters was presented for either a 15 or 30 second stimulus duration followed by a silent interval of equal duration. In a typical experiment, a sequence of such pairs of intervals were presented with the stimulus frequency held fixed and the stimulus level increased in 10 dB steps for successive presentations of the tone. The level in these experiments never exceeded 80 dB SPL. If time allowed, other stimulus frequencies were chosen and the stimulus pattern repeated.

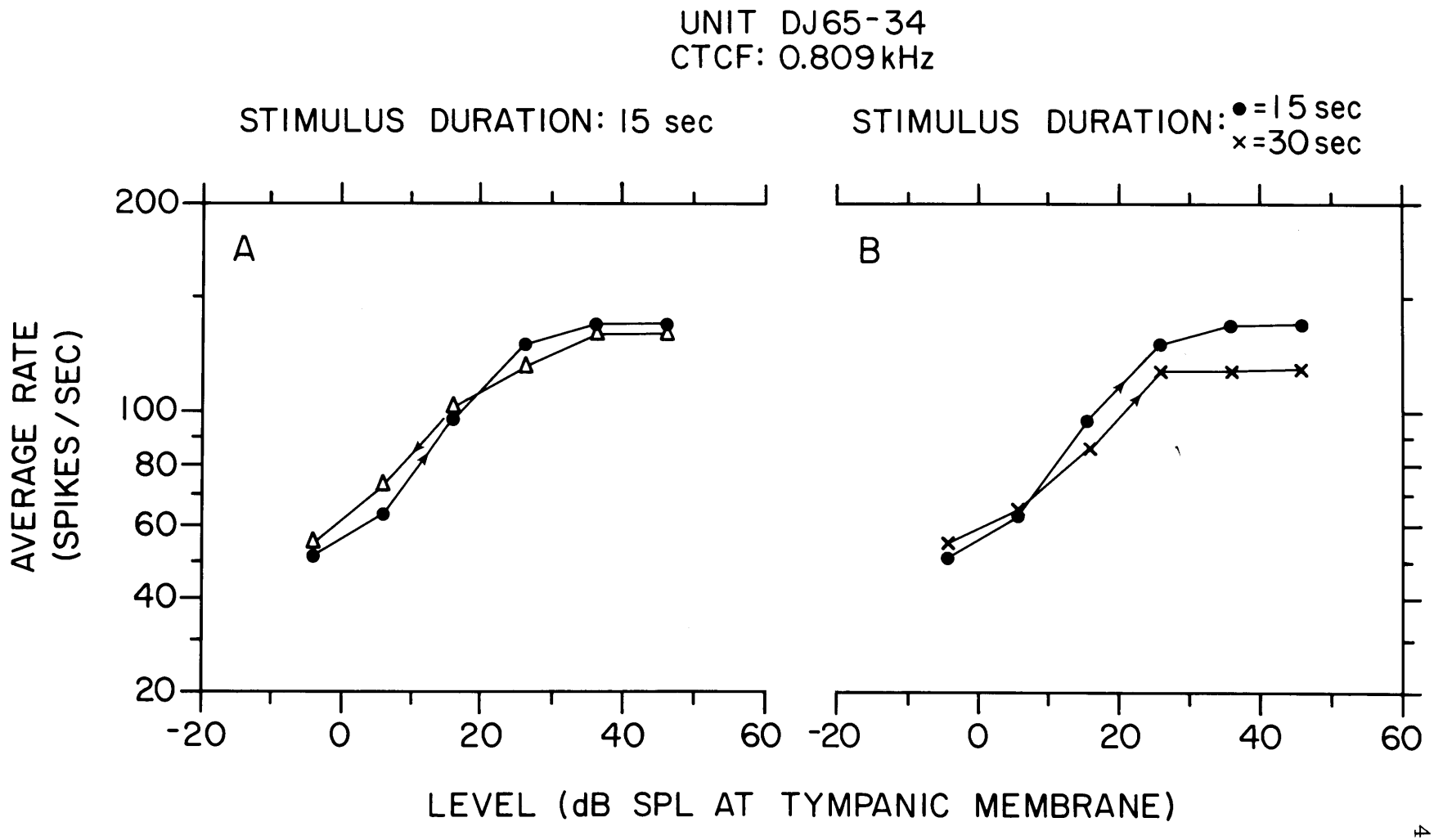
This stimulus protocol reflects an attempt to reduce the effects of previous stimuli upon the response characteristics of auditory-nerve fibers that are known to occur (Kiang et al., 1965: 73, 78-79). In

order to assess how well these effects are reduced, the following data are presented.

Figure 3.2 illustrates average-rate measurements made with the standard stimulus protocol (i.e., the stimulus level is increased from low to high sound pressure levels). Measurements obtained by decreasing the stimulus level from high to low levels are depicted in Figure 3.2a. These data indicate that the direction in which the level is changed does not affect the results with a stimulus duration of 15 seconds. This result may not hold if the stimulus duration is lengthened (30 seconds in this case, Figure 3.2b). The maximum average rates differ but the form of the relationship between average rate and stimulus level is not drastically affected. This result indicates that the average rate of response of auditory-nerve units can change with longer presentation times.

Figure 3.2. Influence of Stimulus Duration and Direction of Level Changes upon Measurements of Average Rate versus Stimulus Level

The data shown in both panels were obtained from one unit; the frequency of the tonal stimulus was the CF of the unit. In each panel, two methods of obtaining average-rate measurements are compared. In panel A, the stimulus level was changed either from low to high levels (circles) or from high to low levels (triangles). The arrows on the curves also indicate the direction of level changes. The stimulus duration was 15 seconds in each case. In panel B, the level was changed from low to high levels while the stimulus duration was either 15 seconds (circles) or 30 seconds (crosses). Each stimulus presentation was followed by a silent interval equal to the stimulus duration; this stimulus protocol applies to both panels.


UNIT DJ65-34
CTCF: 0.809 kHz
STIMULUS DURATION: 15 sec
STIMULUS DURATION: ● = 15 sec
× = 30 sec
A
B
AVERAGE RATE
(SPIKES / SEC)
200
100
80
60
40
20
-20
0
20
40
60
LEVEL (dB SPL AT TYMPANIC MEMBRANE)

3.5 An Analysis of the Computation of PST Histograms

A substantial portion of the measurements of auditory-nerve unit response characteristics are obtained from PST histograms. In order to understand the meaning and the limitations of these measurements, an analysis of the procedure used in computing a PST histogram from recorded auditory-nerve spike data will be presented. Given a train of spike discharges, the computation of a PST histogram consists of three stages:

1. A pulse is triggered from each spike. This pulse serves to denote the time of occurrence of the spike.
2. The time axis is quantized into small intervals of time (bins). The contents of a bin are incremented if a pulse occurs within the interval of time assigned to the bin.
3. The number of spikes which occur in bins at a certain time relative to a series of stimulus markers are added together.

In this manner, the PST histogram reflects a profile of the preferred times of discharge relative to some aspect of the stimulus (usually the onset of the stimulus or a specified phase of a tonal stimulus).

The discharge pattern of an auditory-nerve fiber can be described mathematically as a regular,* stochastic point process. This

*A regular point process cannot have two events (spikes) occurring at exactly the same time. The assumption of regularity is well-founded for auditory-nerve fiber spike trains as seen in interval histograms (Kiang et al., 1965: 94-101).

description assumes that the important aspect of a spike train is the timing of the spikes rather than the waveform of the spike train. The timing timing of events is determined by r(t), the rate of discharge. The function r(t) is defined by the equations:

$$\text{As } \Delta t \to 0: \quad \begin{aligned} &\text{Prob}\{\text{spike occurs in the interval } [t, t+\Delta t)\} = r(t)\,\Delta t \\ &\text{Prob}\{\text{more than one spike occurs in } [t, t+\Delta t)\} = O(\Delta t) \end{aligned} \tag{3.1}$$

where:

$O(\Delta t)$ denotes a function of Δt which approaches zero faster than Δt. In other words,

$$\lim_{\Delta t \to 0} \frac{O(\Delta t)}{\Delta t} = 0.$$

A PST histogram is a statistical estimate of the rate of discharge r(t). The following analysis is concentrated, therefore, upon how each stage in the computation of a PST histogram affects the measurement of r(t).

3.5.1 The triggering of pulses from recorded spikes

If the electrical recording of a spike train were noise-free, the time between some point on the spike waveform (e.g., the peak) and a pulse triggered from the spike would be a constant. Consequently, the onset of the pulse would denote the time of occurrence of the spike. However, recorded spike trains contain some baseline noise. This noise can be due to many factors: physiological noise, electrical

crosstalk from the stimulus, electrode noise, amplifier noise, etc. The presence of baseline noise affects the time at which pulses are triggered. In the following, an analysis of these effects is presented. The analysis leading to equation 3.10 is found in an unpublished work by Dr. Edwin C. Moxon and is presented here with his kind permission.

Let s(t) denote the waveform of one spike and let n(t) denote additive baseline noise. A pulse is triggered at the moment when an electrical signal crosses a trigger level v_T in the proper direction. If no noise were present, the trigger time t_o is given by

$$v_T = s(t_o) \tag{3.2}$$

If noise were present, the pulse would be triggered at the moment $\tilde{t}_o$ such that

$$v_T = s(t_o) + n(\tilde{t}_o) \tag{3.3}$$

In order to express $\tilde{t}_o$ in terms of t_o, we approximate $s(\tilde{t}_o)$ by the first terms of a Taylor series expansion about t_o

$$s(\tilde{t}_o) = s(t_o) + \dot{s}(t_o)(\tilde{t}_o - t_o) \tag{3.4}$$

where

$\dot{s}(t_o)$ denotes the derivative of s(t) at t_o.

The validity of this approximation rests upon the assumption that $\tilde{t}_o - t_o$ is small for the cases of interest. This assumption implies that

the noise n(t) is small in amplitude so that no false triggers or missed triggers occur (i.e., no pulses are triggered from the noise alone or the spike waveform is not reduced below the trigger level v_T). Combining equations 3.2 and 3.3, we have

$$s(\tilde{t}_o) - s(t_o) = -n(\tilde{t}_o). \tag{3.5}$$

Defining m_s to be the slope of the spike at the trigger level ($m_s = \dot{s}(t_o)$) and substituting equation 3.4 into equation 3.5, one obtains

$$m_s \cdot (\tilde{t}_o - t_o) = -n(\tilde{t}_o)$$

or

$$\tilde{t}_o = t_o - \frac{n(\tilde{t}_o)}{m_s} \tag{3.6}$$

The term $n(\tilde{t}_o)/m_s$ represents the jitter in the trigger time due to the presence of baseline noise.

The probability of a spike occurring in the interval Δt must be equal to the probability of a pulse occurring in the corresponding interval $\Delta\tilde{t}$. This statement is equivalent to assuming that no extra pulses are generated. Letting $\tilde{r}(t)$ denote the rate of discharge of the pulse train triggered from the noisy spike train, require

$$r(t)\ \Delta t = \tilde{r}(\tilde{t})\ \Delta\tilde{t} \tag{3.7}$$

Rearranging and passing to the limit of small Δt, equation 3.7 is written

$$\tilde{r}(\tilde{t}) = r(t)\,\frac{dt}{d\tilde{t}} \tag{3.8}$$

From equation 3.6, the derivative is given by

$$\frac{dt}{d\tilde{t}} = 1 + \frac{\dot{n}(\tilde{t})}{m_s} \tag{3.9}$$

where

$\dot{n}(\tilde{t})$ denotes the derivative of the noise n(t) at the time $\tilde{t}$.

Using the equation 3.6 and 3.9, equation 3.8 is written

$$\tilde{r}(\tilde{t}) = r\left(\tilde{t} + \frac{n(\tilde{t})}{m_s}\right) \cdot \left[1 + \frac{\dot{n}(\tilde{t})}{m_s}\right]$$

or

$$\tilde{r}(t) = r\left(t + \frac{n(t)}{m_s}\right) \cdot \left[1 + \frac{\dot{n}(t)}{m_s}\right] \tag{3.10}$$

where

r(t) is the rate of discharge of the spike train.

$\tilde{r}(t)$ is the rate of discharge of the pulse train triggered from the noisy spike train.

m_s is the slope of the spike at the trigger point.

n(t) is the baseline noise.

The rate of discharge $\tilde{r}(t)$ of pulses triggered from a noisy spike train depends not only on r(t), but also on the characteristics of the noise n(t). Two types of errors can occur in estimating r(t) from triggered pulses

1. A time jitter in r(t) is reflected in $\tilde{r}(t)$. The amplitude of this jitter is $n(t)/m_s$.
2. A modulation of r(t) by a term dependent upon the baseline noise is found in the expression for $\tilde{r}(t)$.

These two types of errors can significantly affect measurements of the rate of discharge r(t). Two examples of such errors are illustrated below.

3.5.1(a) Influence of sinusoidal baseline noise synchronized with the acoustic stimulus

Assume the rate of discharge of a spike train to be a constant ($r(t) = r_o$). In this case, $\tilde{r}(t)$ is written

$$\tilde{r}(t) = r_o \cdot \left[1 + \frac{\dot{n}(t)}{m_s} \right] \tag{3.11}$$

Suppose $n(t) = A \cos 2\pi ft$. This situation can occur, for example, if the micropipette records cochlear potentials or electrical artifacts synchronized to the stimulus tone. In this case, equation 3.11 becomes

$$\tilde{r}(t) = r_o \cdot \left[1 + \frac{2\pi fA}{m_s} \cos 2\pi ft \right] \tag{3.12}$$

The estimated rate of discharge of the fiber is synchronized to the sinusoidal baseline noise even though the discharge rate of the fiber is constant. The magnitude of this effect will be discussed after the next example.

3.5.1(b) Influence of random baseline noise

Assume that the rate of discharge r(t) of the spike train is synchronized to a tone so that

$$r(t) = \bar{r} \cdot [1 + 2S_f \cos 2\pi ft] \tag{3.13}$$

where

$\bar{r}$ is the average rate of discharge.

S_f is the amplitude of the synchronized component (i.e., the synchronization index which is discussed in section 2.6.2).

Further assume that n(t) is a stationary stochastic process having a Gaussian amplitude-density with zero mean, variance σ_n^2 and autocorrelation function $R_n(\tau)$. From equation 3.10, the rate of discharge of the triggered pulse train is

$$\tilde{r}(t) = \left(1 + \frac{\dot{n}(t)}{m_s}\right) \cdot \bar{r} \cdot \left[1 + 2S_f \cos 2\pi f\left(t + \frac{n(t)}{m_s}\right)\right] \tag{3.14}$$

Now $\tilde{r}(t)$ is itself a stochastic process; the expected value of $\tilde{r}(t)$ is computed in order to compare its average properties with r(t). This comparison will give a measure of the effects of random baseline noise

on the estimation of r(t). Performing the computation of the expected value of $\tilde{r}(t)$ (Appendix I), one obtains

$$E[\tilde{r}(t)] = \bar{r} \cdot \left[1 + 2S_f \cdot \exp\left\{ -\frac{1}{2} \left(\frac{2\pi f \sigma_n}{m_s} \right)^2 \right\} \cos 2\pi f t \right] \qquad (3.15)$$

The amplitude of the time-varying component is reduced by the factor $\exp\left\{ -\frac{1}{2} \left(\frac{2\pi f \sigma_n}{m_s} \right)^2 \right\}$. This result may be generalized to expressions for r(t) other than given by equation 3.13. If r(t) can be expressed as a finite sum of periodic components, each of which can be written as a Fourier series, the effect of the random baseline noise upon r(t) (as reflected in $\tilde{r}(t)$) is to pass r(t) through the lowpass filter $H_B(f)$ having a Gaussian frequency response. More specifically, if $\tilde{R}(f)$ and R(f) denote the Fourier transforms of $E[\tilde{r}(t)]$ and r(t) respectively, then

$$\tilde{R}(f) = H_B(f) \cdot R(f) \qquad (3.16)$$

where

$$H_B(f) = \exp\left\{ -\frac{1}{2} \left(\frac{2\pi f \sigma_n}{m_s} \right)^2 \right\}$$

Consequently, the amplitudes of the higher-frequency components in r(t), the rate of discharge of the spike train, are reduced in $\tilde{r}(t)$, the rate of discharge of the pulse train triggered from the noisy spike train.

Two simple examples have been considered in which baseline noise affects the measurement of r(t). In each case, the magnitude of the

error introduced by the baseline noise depends upon the ratio of the noise amplitude to the slope of the spike waveform at the triggering level (A/m_s in equation 3.12 and σ_n/m_s in equation 3.15). We approximate m_s by

$$m_s = \frac{s(t_p)}{t_p} \tag{3.17}$$

where

t_p is the time from the baseline to the peak of the spike s(t).

$s(t_p)$ is the peak amplitude of the spike.

Note that this expression is a lower bound upon the slope m_s. The ratios of the noise amplitude to the slope of the spike may be written

$$\frac{A}{m_s} \approx \frac{A}{s(t_p)} \cdot t_p \qquad \frac{\sigma_n}{m_s} \approx \frac{\sigma_n}{s(t_p)} \cdot t_p$$

or

$$\frac{A}{m_s} = \gamma t_p \qquad \frac{\sigma_n}{m_s} = \gamma t_p \tag{3.18}$$

where

γ is the noise-to-spike amplitude ratio $\frac{A}{s(t_p)}$ or $\frac{\sigma_n}{s(t_p)}$.

The parameter γ depends upon the relative amplitudes of the spike and the baseline noise. If the spikes are small in amplitude or if the noise level is large, γ is large. In practical situations, it is the amplitude of the spike which varies from unit to unit and changes slowly with time while recording from one unit. Figure 3.3 depicts spike waveform recorded from two single units. The time-to-peak values of t_p measured from these data is approximately 200 μsec. Equation 3.12 and 3.16 are therefore written

$$\tilde{r}(t) = r_o \cdot [1 + (4\pi\gamma f/10) \cdot \cos 2\pi(f \times 10^3)t] \tag{3.19a}$$

$$H_B(f) = \exp\left\{-\frac{1}{2}(4\pi\gamma f/10)^2\right\} \tag{3.19b}$$

where

f is expressed in kHz.

The values of the noise-to-spike amplitude ratio γ range from 0.04 for "clean" spikes to 0.2 for "noisy" spikes (Figure 3.3). Figure 3.4 depicts the amplitude functions of equations 3.19a and 3.19b. Even for small values of γ, the amount of synchronization in $\tilde{r}(t)$ introduced by a sinusoidal baseline signal can be significant, especially at high frequencies. However, the effect of random baseline noise upon measurements of the synchronization index is not severely affected at high frequencies. For instance, with "clean" spikes ($\gamma = 0.04$), measurements of stimulus-synchronized activity can be made for frequencies up to 10 kHz with a decrease in the synchronization index of

Figure 3.3. Two Examples of Spike Waveforms Recorded from a Single Unit

The left panel depicts a spike waveform recorded relatively free of random baseline noise while the right panel illustrates a spike waveform recorded with a noisier baseline. The parameter γ is the ratio of the peak noise level to the magnitude of the peak amplitude of the spike. The vertical scales are expressed in arbitrary amplitude units (an absolute calibration of spike amplitudes was not made). The horizontal scales denote time relative to the instant a pulse was triggered from each recorded spike. Note that these waveforms represent recorded spikes rather than the actual waveforms of the spikes propagating in the fiber.

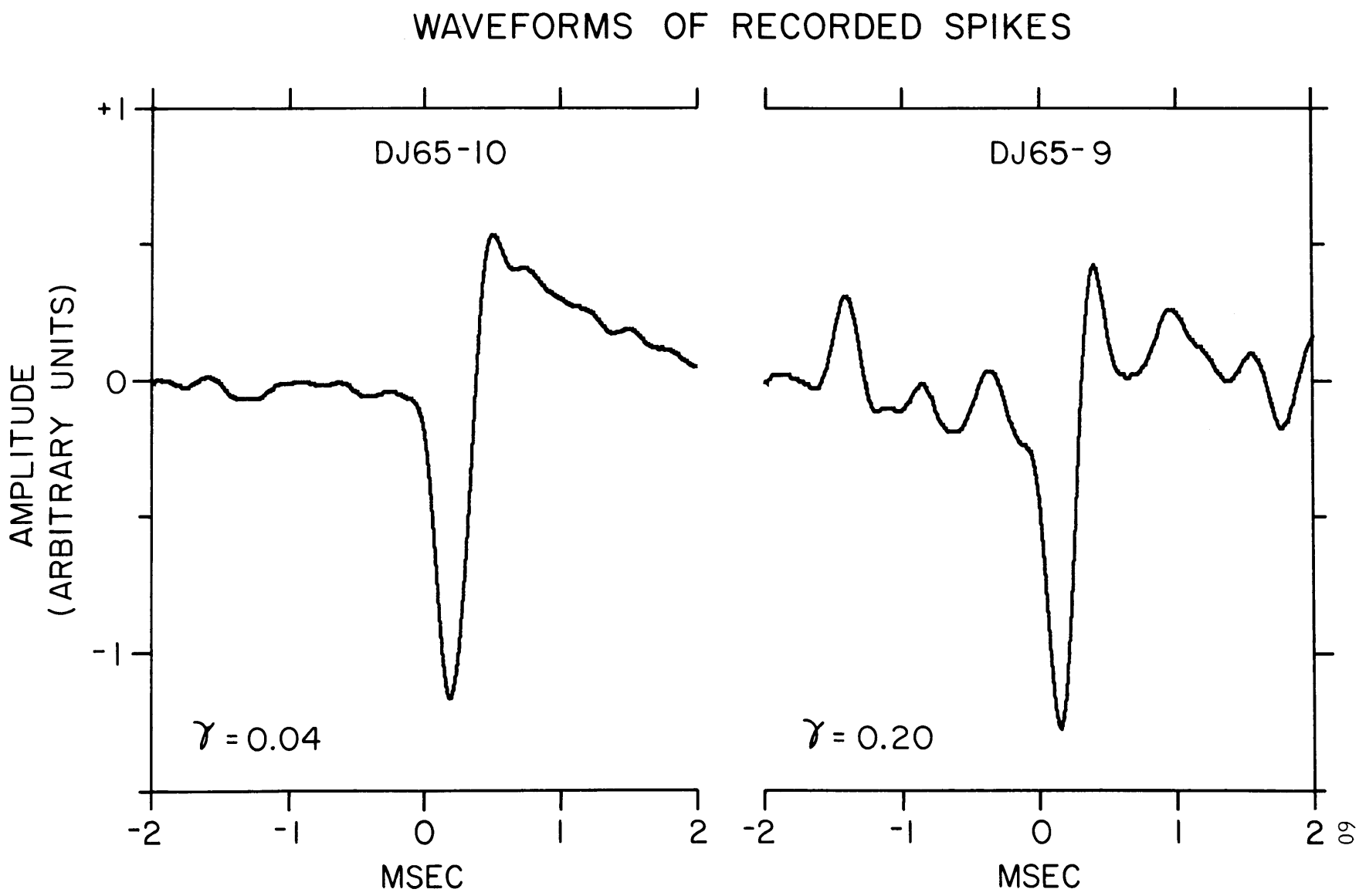
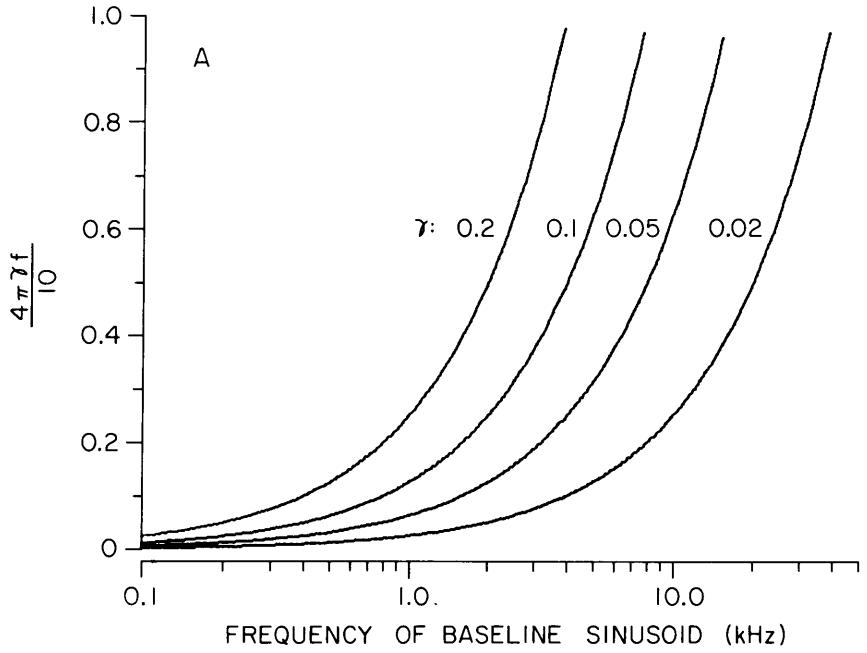

WAVEFORMS OF RECORDED SPIKES
DJ65-10
DJ65-9
AMPLITUDE
(ARBITRARY UNITS)
+1
0
-1
γ = 0.04
γ = 0.20
-2
-1
0
1
2
MSEC
-2
-1
0
1
2
MSEC

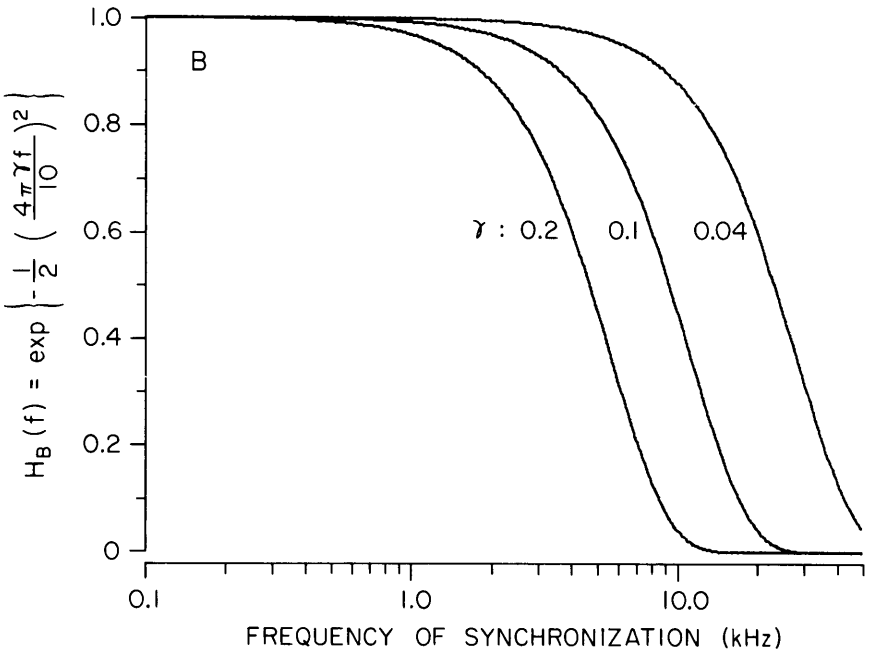
Figure 3.4. Effects of Baseline Signals upon Measurements of Synchronized Activity

In panel A, the spikes are assumed to be recorded in the presence of a sinusoidal baseline signal. The ratio of the peak amplitude of the baseline signal to magnitude of the peak of the spike is denoted by the parameter γ. The spikes are assumed to have a constant rate of discharge. The expression $\frac{4\pi\gamma f}{10}$ is the amplitude of the measured rate of discharge synchronized to the baseline sinusoid.

In panel B, the spikes are assumed to be recorded in the presence of random, Gaussian baseline noise. Here γ is the ratio of the rms amplitude of the noise to the magnitude of the peak of the spike. The spikes are assumed to be synchronized to a sinusoidal stimulus. The transfer function $H_B(f)$ represents the reduction in the measured value of the synchronized response of the fiber due to the effect of the random baseline noise on the triggering process. This reduction depends upon both γ and f, the frequency of the synchronized response.

The expressions derived for these relationships assumed that the time from baseline to the peak of the spike was 200 μsec. The frequency f is expressed in kHz.


A
1.0
0.8
0.6
0.4
0.2
0
$\frac{4\pi\gamma f}{10}$
γ: 0.2
0.1
0.05
0.02
0.1
1.0
10.0
FREQUENCY OF BASELINE SINUSOID (kHz)


B
1.0
0.8
0.6
0.4
0.2
0
$H_B(f) = \exp\left\{-\frac{1}{2}\left(\frac{4\pi\gamma f}{10}\right)^2\right\}$
γ : 0.2
0.1
0.04
0.1
1.0
10.0
FREQUENCY OF SYNCHRONIZATION (kHz)

less than 10%. With "noisy" spikes, however, a significant reduction in the synchronized component can occur at lower frequencies. For example, $H_B(f)$ equals 0.5 at the frequency 4.7 kHz with $\gamma = 0.2$.

3.5.2 Time-axis quantization

The next step in the estimation of r(t) is the quantization of the time axis into intervals of equal duration (bins). The time-quantized pulse train is represented by the sequence x_m where

$$x_m = \begin{cases} 1 & \text{if a pulse occurs in the interval } [m\Delta, (m+1)\Delta) \\ 0 & \text{if no pulse occurs in the interval } [m\Delta, (m+1)\Delta) \end{cases} \tag{3.20}$$

The probability of two pulses occurring in the same bin is assumed to be negligible. The probability of x_m assuming the values 0 or 1 are to be related to $\tilde{r}(t)$, the rate of discharge of the pulse train. With this probability, the expected value of x_m can be computed, thereby allowing an analysis of the effect of time-axis quantization upon measurements of the rate of discharge.

Partition each bin, having duration Δ, into many smaller intervals of duration Δt. As the spike train is assumed to be a <u>regular</u> point process, equation 3.1 applies to the pulse train.

$$\text{Prob}\{\text{pulse occurs in } [t, t+\Delta t)\} = \tilde{r}(t)\,\Delta t$$

$$\text{Prob}\{\text{more than one pulse occurring in } [t, t+\Delta t)\} = O(\Delta t) \tag{3.21}$$

Given that the interval $[t+n\Delta t, t+(n+1)\Delta t)$ lies within a bin of duration Δ and that no more than one pulse can occur in the bin, the probability of a pulse in the m^{th} bin (i.e., the probability $x_m = 1$) is given by

$$\text{Prob}\,[x_m = 1] = \sum_n \tilde{r}(m\Delta + n\Delta t)\Delta t \tag{3.22}$$

Passing to the limit as Δt approaches zero and the number of intervals becomes large, this sum becomes

$$\text{Prob}\,[x_m = 1] = \int_{m\Delta}^{(m+1)\Delta} \tilde{r}(\alpha)\, d\alpha \qquad m = 0, \ldots \tag{3.23}$$

Therefore, the expected value of the contents of the m^{th} bin is written

$$E[x_m] = \int_{m\Delta}^{(m+1)\Delta} \tilde{r}(\alpha)\, d\alpha \qquad m = 0, \ldots \tag{3.24}$$

The estimate $\hat{\tilde{r}}(m\Delta)$ of the rate of discharge at the time $m\Delta$ from the sequence x_m is found by dividing the contents of each bin by the duration of the bin (Δ). The expected value of this estimate $\hat{\tilde{r}}(m\Delta)$ is

$$E[\hat{\tilde{r}}(m\Delta)] = \frac{1}{\Delta} \int_{m\Delta}^{(m+1)\Delta} \tilde{r}(\alpha)\, d\alpha \qquad m = 0, \ldots \tag{3.25}$$

The expected value of the rate of the time quantized pulse train is the pulse-train rate of discharge $\tilde{r}(t)$ passed through a finite-time

integrator having integration time Δ. The average behavior of the sequence x_m is therefore described as the sampling of the output of this integrator every Δ. Letting $h_\Delta(t)$ represent the impulse response of the finite-time integrator, the expected value of x_m is written

$$E[x_m] = \Delta[\tilde{r}(t) \otimes h_\Delta(t)]\Big|_{t=(m+1)\Delta} \qquad m = 0, \ldots \tag{3.26}$$

where

$\tilde{r}(t)$ is the rate of discharge of the pulse train.

$h_\Delta(t)$ is the impulse response of a finite-time integrator specified as:

$$h_\Delta(t) = \begin{cases} \frac{1}{\Delta} & 0 \leq t \leq \Delta \\ 0 & \text{elsewhere} \end{cases} \tag{3.27}$$

$\otimes$ denotes linear convolution.

x_m is the sequence denoting the contents of the m^{th} bin.

The frequency response $H_\Delta(f)$ of the finite-time integrator $h_\Delta(t)$ is given by

$$H_\Delta(f) = \exp\{-j\pi f\Delta\} \cdot \frac{\sin(\pi f\Delta)}{\pi f\Delta} \tag{3.28}$$

Consequently, the rate of discharge $\tilde{r}(t)$ is passed through a lowpass filter $H_\Delta(f)$ whose magnitude $|H_\Delta(f)| \leq \frac{1}{\pi f\Delta}$; i.e., the magnitude

decreases as frequency increases (see Figure 3.5).

Two sources of errors are introduced in the time-quantization of the pulse train triggered from the recorded spikes: 1) the rate of discharge is passed through a lowpass filter (equation 3.28) and 2) the output of this filter is sampled (equation 3.26). Considering the lowpass filter $H_{\Delta}(f)$, the important parameter of its frequency response is the product of the frequency variable f and the binwidth Δ. If the frequency components in $\tilde{r}(t)$ have periods comparable to the binwidth Δ, their amplitudes can be reduced significantly (Figure 3.5). For example, assume a period histogram is computed with 10 bins per stimulus period. The amplitude of the first harmonic in the rate of discharge is reduced about 1.6%, a very small amount. However, if the rate of discharge has significant components at the second and third harmonics, then amplitudes of these components are reduced by 6.5% and 14.2% respectively. Therefore significant reductions in the amplitudes of these components can occur.

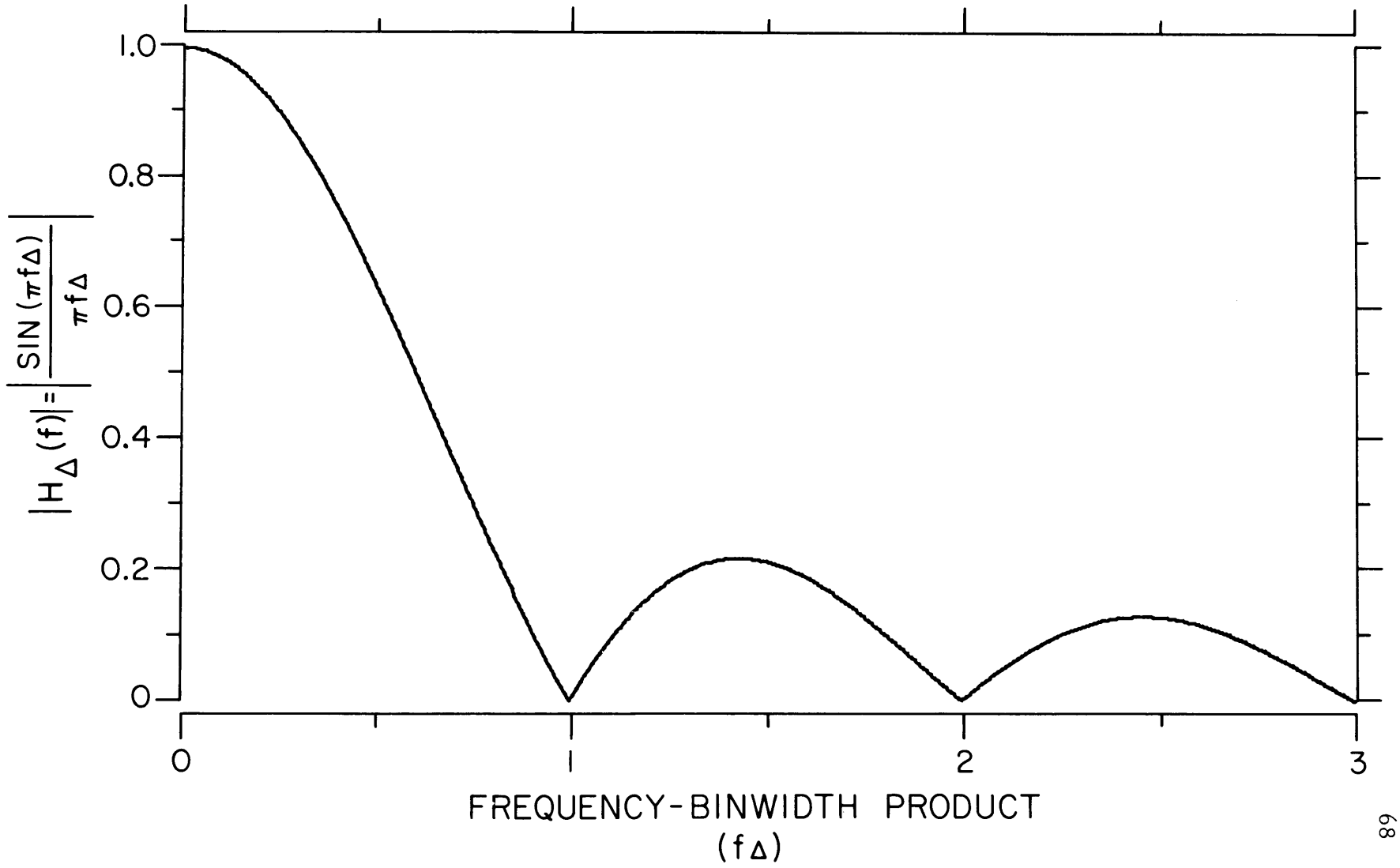
The second possible source of error is the sampling of the output of the finite-time integrator every Δ. If the output, given by $[\tilde{r}(t) \otimes h_{\Delta}(t)] \cdot \Delta$ is bandlimited so that the highest frequency component at f_M satisfies

$$f_M \leq \frac{1}{2\Delta}, \tag{3.29}$$

then all of the information in $[\tilde{r}(t) \otimes h_{\Delta}(t)] \cdot \Delta$ is preserved in its samples x_m. If the largest frequency f_M does not satisfy this

Figure 3.5. Effect of Time Quantization Upon Measurements of the Rate of Discharge

The time of occurrence of a spike is quantized into time intervals (bins) whose duration is Δ seconds. $H_{\Delta}(f)$ represents the transfer function relating the time-quantized rate of discharge to the actual rate of discharge. The magnitude of this transfer function is depicted as a function of $f\Delta$, the product of the frequency (Hz) of a component in the rate of discharge and the binwidth (sec) of the time-quantization. $|H_{\Delta}(f)|$ has zeroes at integer values of $f\Delta$ and the envelope of the peaks decreases with $(f\Delta)^{-1}$.


1.0
0.8
0.6
0.4
0.2
0
$|H_\Delta(f)| = \left|\frac{\text{SIN}(\pi f \Delta)}{\pi f \Delta}\right|$
0
1
2
3
FREQUENCY-BINWIDTH PRODUCT
$(f\Delta)$

inequality, errors can occur. The magnitude of the frequency response $H_\Delta(f)$ at the frequency $\frac{1}{2\Delta}$ is not small and it does decrease rapidly for frequencies larger than $\frac{1}{2\Delta}$.

$$\left|H_\Delta\left(\frac{1}{2\Delta}\right)\right| \equiv \frac{2}{\pi} = 0.637 \tag{3.30}$$

Consequently, the requirement that the output of the finite-time integrator be bandlimited to frequencies less than $\frac{1}{2\Delta}$ is equivalent to requiring the rate of discharge $\tilde{r}(t)$ to be bandlimited. If $\tilde{r}(t)$ is not bandlimited, then components at frequencies larger than $\frac{1}{2\Delta}$ will be reflected as lower frequency components in the sequence x_m. This type of error is termed aliasing.* For instance, if the binwidth Δ is 1 msec $\left(\frac{1}{2\Delta} = 500 \text{ Hz}\right)$, a frequency component in $\tilde{r}(t)$ at 750 Hz will appear in x_m as a 250 Hz component.

In summary, the time-quantization of pulses is mathematically equivalent to passing the rate of discharge $\tilde{r}(t)$ through the system indicated in equation 3.26. Two types of errors can occur in this process:

1. Higher-frequency components in $\tilde{r}(t)$ may be attenuated by the filter $H_\Delta(f)$.
2. Components in $\tilde{r}(t)$ at frequencies larger than $\frac{1}{2\Delta}$ will be found in x_m as lower-frequency components.

---

*For a more complete discussion of sampling and aliasing, see Oppenheim and Schafer (1975, Chapter 1).

The size of these effects depends upon the relationship between Δ, the binwidth, and f, the frequency of a component in $\tilde{r}(t)$.

3.5.3 The computation of the PST histogram

Once the pulse train is quantized in time, the contents of bins occurring at the same time relative to a set of stimulus markers are added together in order to compute a PST histogram. Let the histogram contain M bins, each of duration Δ. The interval between stimulus markers is therefore $M\Delta$. Denoting the contents of the m^{th} bin in the histogram by g_m, the computation of a PST histogram is described by

$$g_m = \sum_{i=0}^{R-1} x_{m+iM} \qquad m = 0, \ldots, M-1 \tag{3.31}$$

where

x_m is the sequence of ones and zeroes resulting from the time-quantization of the pulse train derived from the recorded spikes. The index m is assumed to vary over the range $0 \leq m < B$.

M is the number of bins in the histogram.

$R = B/M$ is the number of stimulus markers used in the computation of the histogram. For simplicity, R is assumed to be an integer.

g_m is the contents of the m^{th} bin of the histogram.

The procedure described by equation 3.31 can be interpreted as passing the sequence x_m through a digital filter whose input/output relationship is characterized by the difference equation given in equation 3.31. The output of this filter is considered only for $m = 0, \ldots, M-1$. The unit-sample response of this filter, $(h_{PST})_m$, is found by setting the sequence x_m equal to the sequence

$$x_m = \delta_m \tag{3.32}$$

where

δ_m is the unit sample. $\delta_m = 1$ for $m = 0$ and $\delta_m = 0$ for $m \neq 0$.

Therefore, the unit-sample response is given by

$$(h_{PST})_m = \sum_{i=0}^{R-1} \delta_{m+iM} \tag{3.33}$$

The Fourier transform $H_{PST}(e^{j\omega})$ of the unit-sample response $(h_{PST})_m$ is the frequency response of this digital filter. $H_{PST}(e^{j\omega})$ is given by

$$H_{PST}(e^{j\omega}) = \sum_{m=-\infty}^{\infty} (h_{PST})_m \, e^{-j\omega m} \tag{3.34}$$

Therefore, substituting equation 3.33 into equation 3.34, one obtains

$$H_{PST}(e^{j\omega}) = \exp\left\{ j \frac{\omega(R-1)\, M}{2} \right\} \cdot \frac{\sin(\omega RM/2)}{\sin(\omega M/2)} \qquad (3.35)$$

where

$H_{PST}(e^{j\omega})$ is the frequency response of the procedure involved in the computation of a PST histogram.

ω is the digital frequency variable. $0 \leq \omega < 2\pi$.

The frequency $\omega = \pi$ corresponds to the maximum frequency allowed in the computation of the sequence x_m ($\frac{1}{2\Delta}$, equation 3.29). Therefore ω is related to the frequency variable f by

$$\omega = 2\pi f\Delta \qquad \text{for} \quad |f| < \frac{1}{2\Delta} \qquad (3.36)$$

In terms of the variable f, equation 3.35 is written

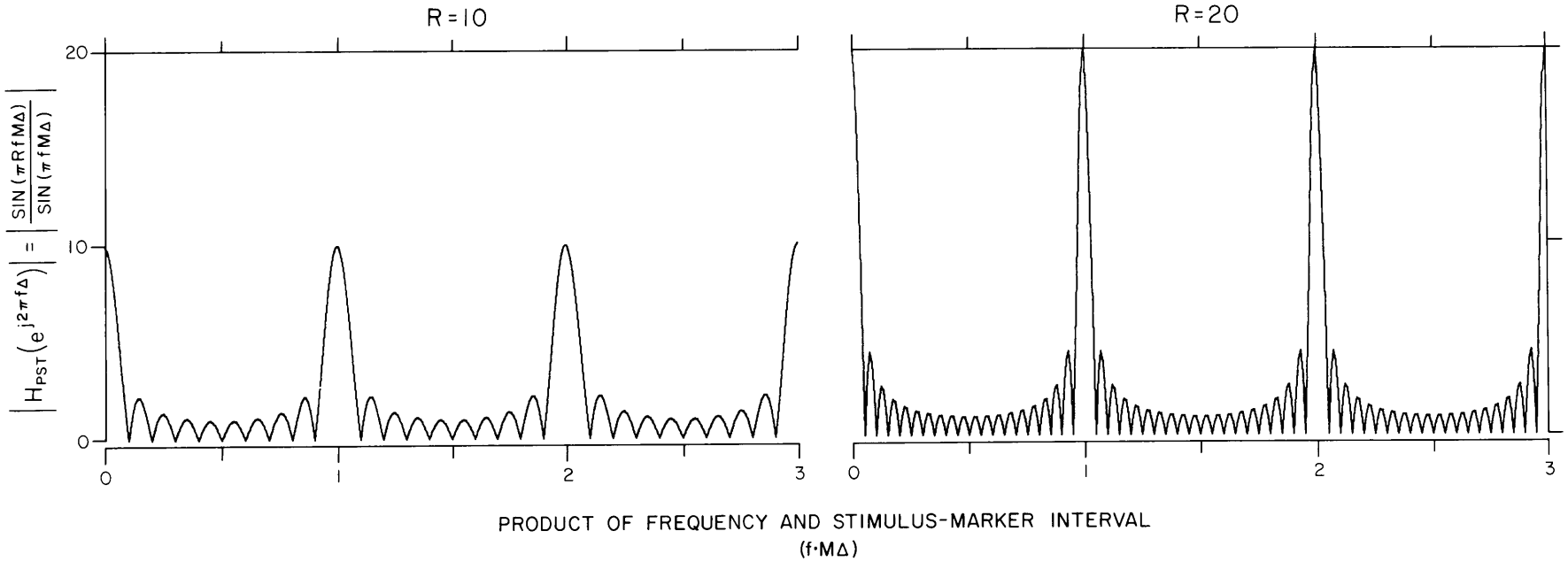
$$H_{PST}(e^{j2\pi f\Delta}) = \exp\{ j\pi fM \cdot (R-1)\Delta \} \cdot \frac{\sin(\pi RfM\Delta)}{\sin(\pi fM\Delta)} \qquad (3.37)$$

The magnitude $|H_{PST}(e^{j2\pi f\Delta})|$ of this frequency response consists of a series of dominant peaks occurring at integer multiples of the frequency $f = \frac{1}{M\Delta}$ (Figure 3.6). The height of these peaks is equal to R, the number of stimulus markers in the interval of measurement. Although the height of these peaks grow with the number of stimulus markers, the magnitude $|H_{PST}(e^{j2\pi f\Delta})|$ is independent of R for frequencies lying between the dominant peaks. Therefore, components

Figure 3.6. Effect of the Computation of a PST Histogram Upon Measurements of the Rate of Discharge

A PST histogram is computed by adding together the contents of bins occurring at the same time relative to a set of stimulus markers. $H_{PST}(e^{j2\pi f\Delta})$ represents the transfer function relating the waveform of the PST histogram to the sequence of time-quantized pulses. The magnitude of this function is depicted as a function of $f \cdot M\Delta$ for two values of R, the number of stimulus markers in the observation interval. The frequency of a component in the rate of discharge of the pulse train is denoted by f and $M\Delta$ is the interval between stimulus markers.

The magnitude of $H_{PST}(e^{j2\pi f\Delta})$ shows large peaks coinciding with integer multiples of $f \cdot M\Delta$; the frequencies at which these peaks occur correspond to harmonics of the synchronizing frequency $1/M\Delta$. As R is increased, the height of each of the large peaks increases while the amplitudes of the smaller peaks remain constant. The amplitude of $|H_{PST}(e^{j2\pi f\Delta})|$ for integer multiples of $f \cdot M\Delta$ is R; the envelope of the smaller peaks ($f \cdot M\Delta$ not equal to an integer) is $|\sin(\pi f M\Delta)|^{-1}$.


R=10
R=20
20
10
0
0
1
2
3
$\left|H_{PST}\left(e^{j2\pi f\Delta}\right)\right| = \left|\frac{\text{SIN}(\pi R f M\Delta)}{\text{SIN}(\pi f M\Delta)}\right|$
PRODUCT OF FREQUENCY AND STIMULUS-MARKER INTERVAL
(f·MΔ)

in the sequence x_m having frequencies that do not coincide with these peaks will tend to be attenuated. For example, if a period histogram is synchronized to a tone of frequency f_o, frequencies harmonic with the stimulus frequency (given by nf_o; $n = 0, 1, \ldots$) are emphasized in the histogram. Inharmonic components (i.e., components whose frequencies f do not equal nf_o) are reduced in amplitude. To see the magnitude of this effect, assume the synchronizing frequency f_o is 1.000 kHz ($M\Delta$ = 1 msec) and that data lasting 30 seconds is used to compute a period histogram. Therefore, the number of stimulus markers R is 3×10^4. Using equation 3.37, a component of frequency 1.001414 kHz is attenuated by a factor of 137 (approximately 43 dB) relative to a component having a frequency exactly equal to 1.000 kHz. Therefore, the only frequency components found in the expected value of a period histogram are those harmonic with the synchronizing frequency $f_o = \frac{1}{M\Delta}$.

3.5.4 Summary

Figure 3.7 summarizes the results of a mathematical model of the steps involved in the computation of a PST histogram. The rate of discharge r(t) is passed through a series of lowpass filters ($H_B(f)$ and $H_\Delta(f)$) before components harmonic to the period of the PST histogram are selected by the filter $H_{PST}(e^{j2\pi f\Delta})$. By using relatively "clean" spikes, the cutoff frequency of the filter $H_B(f)$ is pushed to higher frequencies. The choice of a PST histogram bin-width, which is much shorter than the period of the highest frequency

Figure 3.7. A Model for the Computation of a PST Histogram

The upper portion of the figure indicates the steps involved in the computation of a PST Histogram. Pulse are first triggered from the recorded spikes. The times of occurrence of these pulses are quantized to intervals of duration Δ, the binwidth of the histogram. Finally, the PST histogram is computed by adding together the contents of bins which occur at the same time relative to a set of stimulus markers.

The lower portion of the figure depicts a block diagram of a mathematical model of the steps involved in the computation of a PST histogram. The rate of discharge of the spike train is denoted by r(t). The rate of discharge of the pulse train triggered from the recorded spikes, which are present in random baseline noise, is given by $\tilde{r}(t)$. The rates of discharge r(t) and $\tilde{r}(t)$ are related by a linear, time-invariant filter having the transfer function $H_B(f)$.

$$H_B(f) = \exp\left\{-\frac{1}{2}\left(\frac{2\pi\sigma_n f}{m_s}\right)^2\right\}$$

where

σ_n is the rms amplitude of the baseline noise.

m_s is the slope of the spike waveform at the triggering level.

(3.7) The time-quantization of the triggered pulses results in the sequence x_m. This sequence consists of the sampled values of the output of a linear, time-invariant filter represented by $H_\Delta(f)$; the samples are taken at the times $(m+1)\Delta$.

$$H_\Delta(f) = \exp\{-j\pi f\Delta\} \cdot \frac{\sin(\pi f\Delta)}{\pi f\Delta}$$

The contents of the m^{th} bin of the histogram, g_m, is related to the sequence of time-quantized pulses x_m by the difference equation

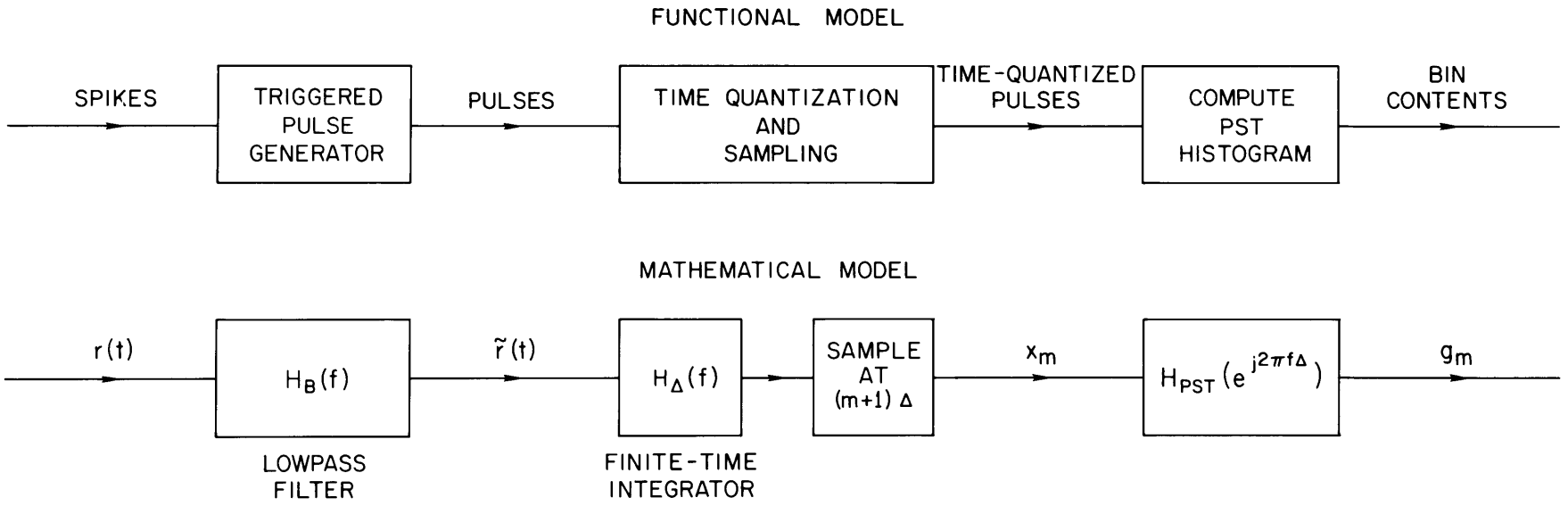
$$g_m = \sum_{i=0}^{R-1} x_{m+iM} \qquad m = 0, \ldots, M-1$$

which has a transfer function $H_{PST}(e^{j2\pi f\Delta})$

$$H_{PST}(e^{j2\pi f\Delta}) = \exp\{j\pi fM \cdot (R-1)\Delta\} \cdot \frac{\sin(\pi RfM\Delta)}{\sin(\pi fM\Delta)}$$

This entire mathematical model describes the <u>average</u> influence of each of the steps in a PST histogram computation upon the rate of discharge of the spike train.

MODEL OF PST HISTOGRAM COMPUTATION



component in r(t), reduces both the effect of $H_{\Delta}(f)$ and the aliasing of the waveform. If these general rules cannot be met, equations 3.19b and 3.28 allow a computation of the degree r(t) may be changed in the computation of a PST histogram. For instance, synchronized responses to a tone can be theoretically measured for stimulus frequencies as large as 10 kHz if the spikes are relatively noise-free and if the binwidth of the histogram is small compared to the period of the stimulus.

3.6 Measures of Response

Two aspects of the response of an auditory-nerve fiber to tonal stimuli are of interest in this thesis: the average rate of response and the synchronized response. In this section, measures of these responses are defined and their statistics analyzed. The statistics of these measures are difficult to compute unless the triggered pulse train is assumed to be described by a Poisson process. The assumption of Poisson statistics is explicitly denoted by the use of the notation $\lambda(t)$ for the rate of discharge. When the results of derivations using the Poisson assumption are to be applied to auditory-nerve fiber response patterns, the notation for the rate of discharge is changed to $r(t)$. Knowledge of the statistics of the underlying point process description of auditory-nerve fiber spike trains is limited to some details concerning spontaneous activity. In this instance, the interval between successive spikes can be described as statistically independent random variables (Rodieck et al., 1962). This is also a characteristic of a Poisson-process. However, a Poisson-process description requires the probability density function of the interval to be an exponential

$$p_\tau(\tau) = \lambda_o e^{-\lambda_o \tau} \tag{3.38}$$

Equation 3.38 approximately describes interval histograms computed from the spontaneous activities of auditory-nerve fibers; there are fewer short intervals than necessary to be consistent with a Poisson assumption.

3.6.1 Average discharge rate

Given the sequence x_m defined in section 3.5.2, the measure of the average rate is defined as

$$\bar{r} = \frac{1}{T} \sum_{m=0}^{B-1} x_m \tag{3.39}$$

where

T is the length of the observation interval.

B is the number of bins, each having duration Δ, in the observation interval. Therefore $B\Delta = T$.

The expected value of $\bar{r}$ is given by

$$E[\bar{r}] = \frac{1}{T} \sum_{m=0}^{B-1} E[x_m] \tag{3.40}$$

Substituting equation 3.23, equation 3.40 becomes

$$E[\bar{r}] = \frac{1}{T} \int_0^T \tilde{r}(\alpha)\, d\alpha \tag{3.41}$$

The expected value of $\bar{r}$ is the long-term average value of the rate of discharge $\tilde{r}(t)$. As the low-frequency components of r(t) are preserved in $\tilde{r}(t)$ (Figure 3.4b), the expected value of $\bar{r}$ is equivalent to the long-term average value of r(t).

In order to compute the variance of $\bar{r}$, Poisson statistics are assumed in order to simplify the derivation. We compute, therefore, the variance of $\bar{\lambda}$. As the random variables x_i and x_j are statistically independent ($i \neq j$) for a Poisson process, the variance of $\bar{\lambda}$, $\sigma^2_{\bar{\lambda}}$, is given by

$$\sigma^2_{\bar{\lambda}} = \frac{1}{T^2} \sum_{m=0}^{B-1} \sigma^2_{x_m} \tag{3.42}$$

where

$\sigma^2_{x_m}$ is the variance of the random variable x_m.

The variance of x_m is found by

$$\sigma^2_{x_m} = \left[\int_{m\Delta}^{(m+1)\Delta} \lambda(a)\, da \right] \cdot \left[1 - \int_{m\Delta}^{(m+1)\Delta} \lambda(a)\, da \right] \tag{3.43}$$

Consequently, $\sigma^2_{\bar{\lambda}}$ is given by

$$\sigma^2_{\bar{\lambda}} = \frac{1}{T^2} \left\{ \int_0^T \lambda(a)\, da - \sum_{m=0}^{B-1} \left[\int_{m\Delta}^{(m+1)\Delta} \lambda(a)\, da \right]^2 \right\} \tag{3.44}$$

This expression cannot be evaluated without an explicit specification of $\lambda(t)$. However, this variance can be bounded by neglecting the second term in equation 3.44

$$\sigma^2_{\bar{\lambda}} < \frac{1}{T^2}\int_0^T \lambda(a)\ da$$

or

$$\sigma^2_{\bar{\lambda}} < \frac{E[\bar{\lambda}]}{T} \tag{3.45}$$

Assuming this result derived for a Poisson process applies to measurements from auditory-nerve spike trains, we obtain

$$E[\bar{r}] = \frac{1}{T}\int_0^T \tilde{r}(a)\ da \tag{3.46a}$$

$$\sigma_{\bar{r}} < \left(\frac{E[\bar{r}]}{T}\right)^{1/2} \tag{3.46b}$$

The range of statistical variability in the measurement of the average rate is indicated in the figures by $\pm 2\left(\frac{E[\bar{r}]}{T}\right)^{1/2}$, an upper bound upon plus or minus two standard deviations of $\bar{r}$.

3.6.2. Synchronization index

A low-frequency tone can evoke a response synchronized to the individual cycles of the stimulus. Discharges tend to occur preferentially in one portion of a cycle. More precisely stated, the rate of discharge r(t) of an auditory-nerve fiber is assumed to be a periodic function, the period equal to the period of the sinusoidal stimulus. Consequently, r(t) may be expressed as a Fourier series

$$r(t) = \bar{r} \cdot \left[1 + 2 \sum_{k=1}^{\infty} S_{kf} \sin 2\pi(kft + \phi_{kf}) \right] \qquad (3.47)$$

where

$\bar{r}$ is the average rate of discharge.

S_{kf} is the amplitude of the k^{th} harmonic of the frequency f in the response.

ϕ_{kf} is the phase of the k^{th} harmonic.

In general, the time-varying component of r(t) is specified completely only if the values of S_{kf} and ϕ_{kf} are given for all values of k. We shall use the parameter S_f, the magnitude of the fundamental component of the response, as a measure of synchrony (i.e., the degree of time-locking in r(t) to a tonal stimulus). Such a measure has been used by other investigators (Goldberg and Brown, 1969; Anderson et al., 1971; Littlefield et al., 1972). Consistent with Anderson et al., the measure S_f is termed the synchronization index. We shall refer to ϕ_f, the phase angle of the fundamental component, as the phase of the response to a tone. These measures will accurately reflect the synchronized response only if the fundamental component dominates the expression for r(t) (see equation 3.47).

If two tones are present in the stimulus, the rate of discharge may contain components synchronized to several inharmonic frequencies f_i. In this case, r(t) is written

$$r(t) = \bar{r} \cdot \left[1 + \sum_{i} 2 \sum_{k=1}^{\infty} S_{kf_i} \sin 2\pi(kf_i t + \phi_{kf_i}) \right] \qquad (3.48)$$

The measure of the synchrony of r(t) to a component of frequency f_i is S_{f_i}; similarly, the phase of this component is expressed by ϕ_{f_i}. Note that the stimulus need not contain a tone at the frequency f_i. For instance, f_i may be the combination frequency $2f_1 - f_2$.

As shown in section 3.5.3, the computation of a period histogram synchronized to the frequency f attenuates any components in r(t) which are inharmonic with the frequency f (Figure 3.6). Consequently, if the rate of discharge r(t) is given by either equation 3.47 or 3.48, the expected value of a period histogram synchronized to the frequency f is given by

$$E[g_m] = R\Delta \cdot \bar{r} \cdot \left[1 + 2 \sum_{k=1}^{\infty} S_{kf} \sin 2\pi \left(\frac{km}{M} + \phi_{kf} \right) \right] \quad m = 0, \ldots, M-1 \qquad (3.49)$$

where

- g_m denotes the contents of the m^{th} bin, $m = 0, \ldots, M-1$ of the period histogram.
- R is the number of stimulus markers in the observation interval.
- Δ is the binwidth of the histogram.

In writing equation 3.49, two assumptions were made: 1) the components synchronized to frequencies inharmonic to f are negligible

and 2) the number of bins in a period is sufficiently large so that

$$\int_{k\Delta}^{(k+1)\Delta} r(\alpha)\, d\alpha \approx r(k\Delta)\Delta \tag{3.50}$$

Note that the estimate of the rate of discharge of the component in the response synchronized to the frequency f is given by $g_m/(R\cdot\Delta)$ The expected value of the period histogram normalized in this manner is given by

$$E[g_m/R\Delta] = \bar{r}\cdot\left[1 + 2\sum_{k=1}^{\infty} S_{kf} \sin 2\pi\left(\frac{km}{M} + \phi_{kf}\right)\right] \tag{3.51}$$

The synchronization index S_f and the phase of the response ϕ_f are given by

$$S_f = \left(S_{f,c}^2 + S_{f,s}^2\right)^{1/2} \tag{3.52a}$$

$$\phi_f = \frac{1}{2\pi}\tan^{-1}\left(\frac{S_{f,c}}{S_{f,s}}\right) \tag{3.52b}$$

where

$$S_{f,c} = \sum_{m=0}^{M-1} \frac{g_m}{N}\cos\frac{2\pi m}{M} \tag{3.53a}$$

$$S_{f,\,s} = \sum_{m=0}^{M-1} \frac{g_m}{N} \sin \frac{2\pi m}{M} \tag{3.53b}$$

$S_{f,\,c}$ and $S_{f,\,s}$ are the quadrature components of the adjusted period histogram g_m/N.

N is the number of spikes in the histogram.

$$N = \sum_{m=0}^{M-1} g_m = MR\bar{r}\Delta. \tag{3.54}$$

The synchronization index and the phase have the following properties

1. S_f assumes values ranging from zero (i.e., no component synchronized to the frequency f is present in r(t).) to one (i.e., complete synchrony; all of the discharges lie in one bin).
2. The expected value of S_f is independent of the average discharge rate $\bar{r}$. As written in equation 3.49 and reflected in equation 3.52a, the synchronization index is normalized with respect to $\bar{r}$.
3. The phase ϕ_f has units of fraction of a period, zero corresponding to 0° and 1 corresponding to 360°. Positive values of ϕ_f correspond to lead.

The expected value and standard deviation of S_f and ϕ_f are difficult to compute analytically. In order to obtain an estimate of these statistics, a simulation of a period histogram computed from

a Poisson process with known values of S_f and ϕ_f was performed. Appendix II details the simulation method. The empirical formula for the standard deviation of S_f derived in Appendix II is

$$\sigma_{S_f} = \frac{\left[\frac{\sqrt{4-\pi}}{2} + 0.25(1 - \exp\{-10\, S_f\})\right] \cdot \left[1 - S_f^2\right]}{\sqrt{N}} \qquad (3.55)$$

where

N is the number of spikes in the period histogram.

For values of S_f greater than about 0.4, the standard deviation of S_f is proportional to $\left(1 - S_f^2\right)$. Where indicated, the error bracket for the measure S_f span $\pm 2\sigma_{S_f}$ where σ_{S_f} is computed according to equation 3.55.

3.6.3 Scales for plotting $\bar{r}$ and S_f

In choosing scales for plotting $\bar{r}$ and S_f versus some parameter (e.g., tone level), it is convenient if statistical variations of these measures span the same distance on the plot regardless of the location of the data point. The standard deviation of $\bar{r}$ increases with larger values of $\bar{r}$; consequently, the average rate should be plotted on a compressing scale (i.e., the scale grows slower than a linear relationship with $\bar{r}$). The standard deviation of S_f decreases as S_f approaches 1; the synchronization index should

therefore be plotted on an entirely different scale than the average rate. Preferably, S_f should be plotted on an expanding scale.

We have chosen to plot average rate on a square-root ordinate scale; distance along the scale is proportional to the square-root or $\bar{r}$. In this manner, the error bracket $\bar{r} \pm 2\sigma_{\bar{r}}$ spans a distance d given by

$$d = c\left(\left(\bar{r}+2\sigma_{\bar{r}}\right)^{1/2} - \left(\bar{r}-2\sigma_{\bar{r}}\right)^{1/2}\right) \tag{3.56}$$

where

c is a constant of proportionality.

The term $\frac{\sigma_{\bar{r}}}{\bar{r}} = \frac{2}{\bar{r}^{1/2}T}$, as derived from equation 3.46b, is usually much less than one (for example, if $\bar{r} = 49$ spikes/sec and T = 30 seconds, $\frac{2}{\bar{r}^{1/2}T} = 0.009$) in which case

$$d = \frac{2c}{T} \tag{3.57}$$

In plotting the average rate on a square-root scale, the distance spanned by the error bracket $\bar{r} \pm 2\sigma_{\bar{r}}$ is <u>independent of $\bar{r}$</u> as long as $\frac{2}{\bar{r}^{1/2}T} \ll 1$.

The synchronization index S_f is plotted so that $1 - S_f$ is placed on a logarithmic scale. As a very rough approximation, the standard deviation of S_f is proportional to $1 - S_f$.

$$\sigma_{S_f} \approx c_1(1-S_f) \tag{3.58}$$

The standard deviation of $1 - S_f$ is identical to the standard deviation of S_f. The distance d spanned by the error brackets when plotting $(1-S_f)$ on a logarithmic scale is given by

$$d = c_2 \cdot \left[\ln (1-S_f+2\sigma_{1-S_f}) - \ln (1-S_f-2\sigma_{1-S_f})\right] \tag{3.59}$$

Using equation 3.58, we have

$$d = c_2 \ln \left[\frac{(1+2c_1)}{(1-2c_1)}\right] \tag{3.60}$$

Therefore in plotting S_f on a scale equivalent to a logarithmic scale for $1 - S_f$, the error brackets span the same distance independent of S_f. However, the constant of proportionality c_1 depends upon the number of spikes in the histogram (equation 3.55). Therefore, the size of the error bracket is independent of the measured value of S_f only if the number of spikes in the histogram is the same for each data point.

3.7 Interval Histograms of Synchronized Response Pattern

The frequencies of all components present in a synchronized discharge pattern are not always known. For instance, the pattern of response to an inharmonic two-tone stimulus may contain many components (e.g., combination frequency $2f_1 - f_2$, sum and difference frequencies). The period histogram will only properly reflect components synchronized with the period of the histogram. Even if all the frequencies of the response components were known, it would take a great deal of time to assess the relative amplitudes of these components by computing period histograms synchronized to each inharmonic component. The interval histogram also reflects periodic components present in the rate of discharge r(t) (Hind et al., 1967). A quantitative estimate of the frequencies of these components and their relative amplitudes by visual inspection of the interval histogram is, at best, difficult since the pattern of peaks and valleys in the histogram can be intricate. However, the Fourier transform of the interval histogram reveals all the spectral components of r(t) and the amplitude of each component in the transform is related to the synchronization index. This section explores using the spectra of interval histograms to discern which frequency components are present in r(t).

3.7.1 Derivation of the spectrum of the interval histogram

We shall derive analytically the spectrum of an interval histogram

computed from an inhomogeneous Poisson process. Letting $\lambda(t)$ re represent the rate of discharge of a Poisson process, the probability density function (pdf) of τ, the interval between an event occurring at time t and the next event, is written

$$p_{\tau|t}(\tau|t) = \lambda(t+\tau) \exp\left\{-\int_t^{t+\tau} \lambda(a)\, da\right\} \tag{3.61}$$

Let $p_\tau(\tau)$ represent the interval density obtained if the time of the event initiating the interval τ is ignored. This pdf is given by

$$p_\tau(\tau) = \lim_{T\to\infty} \frac{\int_{-T/2}^{T/2} \lambda(t)\,\lambda(t+\tau) \exp\left\{-\int_t^{t+\tau} \lambda(a)\, da\right\} dt}{\int_{-T/2}^{T/2} \lambda(t)\, dt} \tag{3.62}$$

The numerator of equation 3.62 is obtained by multiplying $p_{\tau|t}(\tau|t)$ (equation 3.61) by the probability of an event occurring at time t ($\lambda(t)dt$) and integrating with respect to t. The denominator provides the normalization so that $p_\tau(\tau)$ has area 1.

Let $\lambda(t)$ be of the form

$$\lambda(t) = \bar{\lambda} \cdot \left[1 + 2 \sum_i S_{f_i} \sin 2\pi(f_i t + \phi_{f_i})\right] \tag{3.63}$$

Assuming that $\frac{2\bar{\lambda} S_{f_i}}{\pi f_i} \ll 1$, the substitution of equation 3.63 into equation 3.62 yields

$$p_\tau(\tau) = \bar{\lambda}\, e^{-\bar{\lambda}\tau} \cdot \left[1 + 2 \sum_i S_{f_i}^2 \cos(2\pi f_i \tau)\right] \qquad (3.64)$$

This equation approximately describes an interval histogram measured from a Poisson process having a rate of discharge given by equation 3.63. In the course of the derivation of this equation, the assumption $\frac{2\bar{\lambda}S_{f_i}}{\pi f_i} \ll 1$ was made. For instance, setting $\bar{\lambda}$ = 50/sec (the average rate), f_i = 500 Hz, and S_{f_i} = 1, this term is numerically equal to 0.06. For frequencies greater than 500 Hz, this term is small.

Let P(f) represent the Fourier transform of $p_\tau(\tau)$. For positive frequencies P(f) is given by

$$P(f) = M_{\bar{\lambda}}(f) + \sum_i S_{f_i}^2 M_{\bar{\lambda}}(f - f_i) \qquad (3.65)$$

where

$M_{\bar{\lambda}}(f)$ represents the Fourier transform of $\lambda\, e^{-\lambda\tau}$.

The Fourier transform of the interval histogram consists of the function $M_{\bar{\lambda}}(f)$ centered at zero frequency and the same function scaled by $S_{f_i}^2$ centered about the component frequencies f_i. The amplitude of the spectrum at each frequency f_i is proportional to the square of the synchronization index at the frequency f_i. To the degree than auditory-nerve fiber discharges can be

characterized by a Poisson process, a similar relationship should hold for the spectra of interval histograms computed from auditory nerve discharges.

3.7.2 Considerations of the computation of an interval histogram

As described in section 3.5.4, the computation of a PST histogram can influence measurements of the rate of discharge. Similar effects occur in the computation of an interval histogram. In particular, the time-quantization of the pulse train triggered from the discharges results in a lowpass characteristic being applied to the spectrum of the interval density of the pulse train (section 3.5.2). The contents of the m^{th} bin of a interval histogram, I_m, are related to the pdf of τ, $p_\tau(\tau)$, by

$$I_m = p_\tau(\tau) \otimes h_\Delta(\tau)\Big|_{\tau=(m+1)\Delta} \tag{3.66}$$

where

$h_\Delta(\tau)$ is the impulse response of the finite time integrator defined by equation 3.27.

Δ is the binwidth of the histogram.

$\otimes$ denotes convolution.

The largest frequency in the Fourier transform of I_m is the frequency $f_M = \frac{1}{2\Delta}$. Consequently, any frequency components in r(t)

having larger frequencies than $\frac{1}{2\Delta}$ are reflected as lower frequency components (i.e., aliasing). Considering frequencies in r(t) smaller than $\frac{1}{2\Delta}$, the frequency response of $h_\Delta(t)$, $H_\Delta(f)$, has significant attenuation for frequencies less than $\frac{1}{2\Delta}$ (see Figure 3.5). Therefore, the relative amplitudes of components in r(t) may not be preserved in the spectrum of the interval histogram I_m. After the Fourier transform calculation is performed, the result is multiplied by $|H_\Delta(f)|^{-1}$ so as to correct for the lowpass-filtering effects concomitant in equation 3.66.

3.7.3 An example

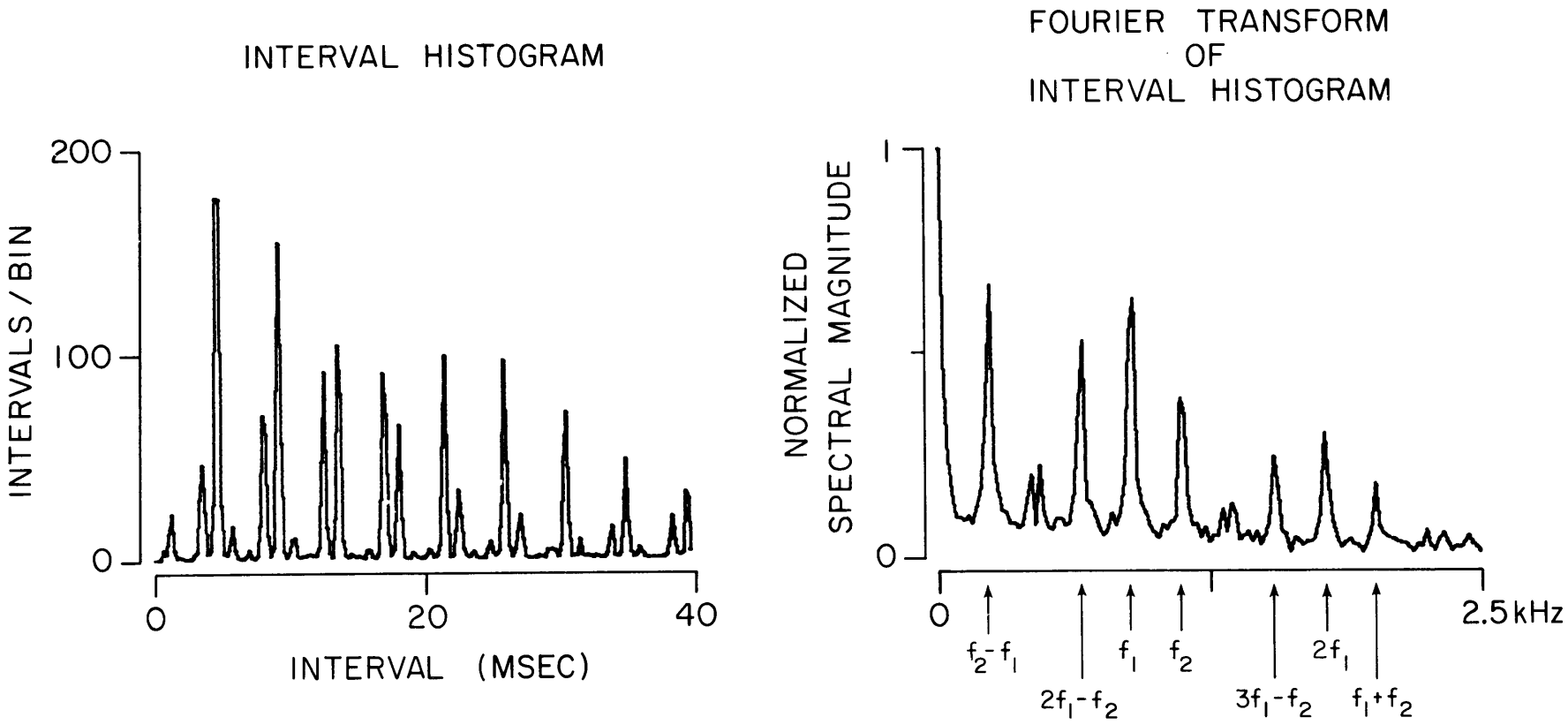
Figure 3.8 depicts the result of calculating the Fourier transform of an interval histogram computed from an auditory-nerve fiber discharge pattern. The stimulus in this instance contained two tones. Many peaks are seen in the transform, each of which presumably corresponds to a synchronized component in the rate of discharge r(t). From this one computation, the frequencies of all of the components in r(t) (having frequencies less than 2.5 kHz) can be determined as well as their relative amplitudes (e.g., the component synchronized to the difference frequency $f_2 - f_1$ is the largest component in the response). A baseline signal is also present in the transform. Presumably, this wideband component is due to the statistical variability present in all interval histograms. The presence of this wideband component limits the accuracy with which measurements of the synchronization index for each component can be made.

Figure 3.8. Example of the Fourier Transform of an Interval Histogram

An interval histogram computed from a two-tone response pattern of an auditory-nerve fiber is shown in the left panel. The stimulus consisted of two tones having frequencies 0.89 kHz (f_1) and 1.12 kHz (f_2). The levels of the tones were 52.2 dB SPL for tone 1 and 52.7 dB for tone 2. This interval histogram was computed from 2700 responses gathered over a one minute interval. The spontaneous discharge rate of the unit was 3.4 spikes/second and its CF was 0.657 kHz.

The Fourier transform of this interval histogram is shown in the right panel. The computational and normalizational procedures are described in section 3.8. Several frequencies related to the stimulus frequencies are indicated along the linear frequency axis.

K451-13
STIMULUS: $f_1 + f_2$



3.8 Computational Methods

Two computers and two computational languages were used in the course of performing the calculations involved in this thesis. A Digital Equipment Corporation PDP-12 computer, operated by the Communications Biophysics Group of the Research Laboratory of Electronics, was used to compute synchronization indices and their phases (equations 3.52a and 3.52b) from period histograms stored on LINC tape. These computations were performed in FOCAL-12, an interpretive language which represents all numbers in floating point with 24-bit mantissas.

Fourier transforms of interval histograms were computed on a PDP-8/E computer operated by the Eaton-Peabody Laboratory. The transforms were computed using the fast Fourier transform algorithm (Oppenheim and Schafer, 1975: Chapter 6) written in OS/8-FORTRAN. This language represents floating point variables with 28-bit mantissas. Interval histograms consisted of 199 bins; each histogram was extended with a sequence of zeroes to form a 256-point sequence. The fast Fourier transform was written to compute a 256-point transform. The magnitudes of these transforms were computed and each was normalized in two ways:

1. The magnitude was scaled by a constant equal to the amplitude of the zero-frequency component. In this way, the amplitude of the component $M_{\lambda}(f)$ is removed from the transform (see equation 3.65).

2. A frequency-dependent normalization was applied to compensate for the lowpass characteristic implicit in equation 3.66. The scaled magnitude was normalized by

$$\left| H_{\Delta}(f) \right|^{-1} = \frac{\pi f \Delta}{\sin (\pi f \Delta)} \qquad 0 \leq f \leq \frac{1}{2\Delta} \tag{3.67}$$

If $P(f)$ represents the Fourier transform of an interval histogram and $P_N(f)$ represents the result of these normalizations, then

$$\left| P_N(f) \right| = \frac{\left| P(f) \right|}{\left| P(0) \right|} \cdot \frac{\pi f \Delta}{\sin (\pi f \Delta)}, \qquad 0 \leq f \leq \frac{1}{2\Delta} \tag{3.68}$$

The frequency $\frac{1}{2\Delta}$ corresponds to the largest frequency computed in the Fourier transform. The first 129 points of each normalized transform magnitude were then stored on LINC tape.

CHAPTER IV

SINGLE-TONE RESPONSE PATTERNS: DATA AND MODEL

The discharge pattern of an auditory-nerve fiber responding to a tone can be specified only if both the synchronized response and the average rate of response are determined. The first portion of this chapter presents measurements of both of these aspects of the response. Data obtained from 169 level series in 111 fibers formed the empirical database of this study. An analytical expression that approximates the rate of discharge is then presented and the dependence of the parameters of this approximation upon the level and frequency of the stimulus is determined. Then we present a model of the responses to single tones applicable for any stimulus frequency and for stimulus levels below 80 dB SPL. The model is intended to summarize how the parameters of the approximation are related to stimulus parameters. In this manner, the response patterns of auditory-nerve fibers to single tones are described in the light of the analytical approximation to the rate of discharge.

4.1 The Waveform of the Rate of Discharge

In this section, the waveform of the synchronized response to single tones is described. In order to characterize the deviation of the waveforms from a sinusoid, we shall define two terms. "Symmetry" indicates symmetry of the waveform with respect to a horizontal line. More specifically, letting r(t) be the rate of discharge and T be the stimulus period, a symmetric waveform satisfies

$$r(t) - r_o = r_o - r\left(t + \frac{T}{2}\right) \tag{4.1}$$

where

r_o is the amplitude of a reference line.

The term "skew" indicates the assymmetry of the waveform with respect to a vertical reference line passing through the mode. A waveform having zero skew satisfies

$$r(t_o+t) = r(t_o-t) \tag{4.2}$$

where

$r(t_o)$ is the mode of the waveform.

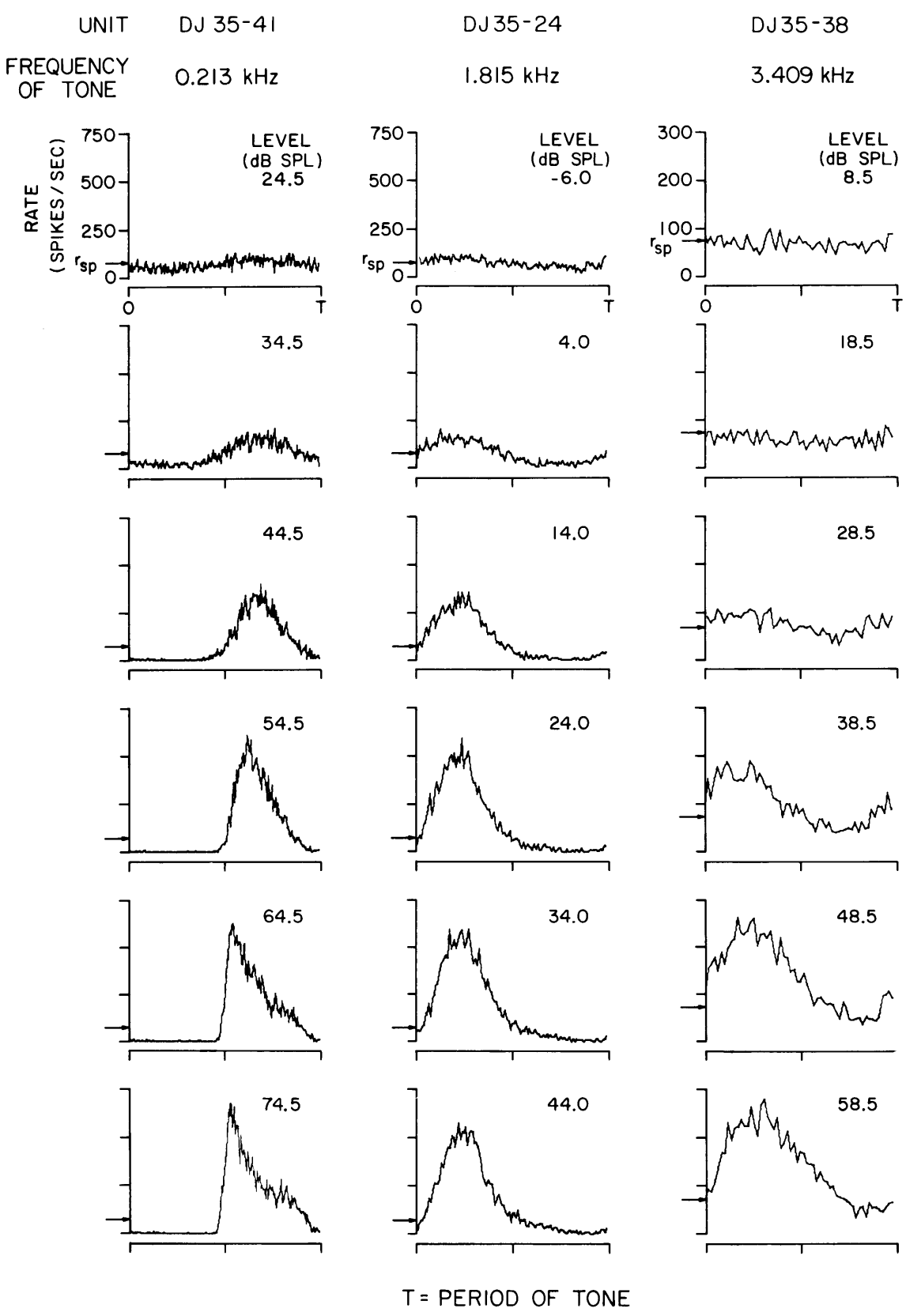
Using this terminology, a sinusoid is symmetric and has no skew.

Figure 4.1 illustrates period histograms obtained from the responses of three auditory-nerve fibers in one cat. These fibers

Figure 4.1. Period Histograms of the Responses to Single Tones

Period histograms synchronized to the positive-going zero-crossing of the oscillator's output are shown for three units obtained from one animal. The horizontal scale represents one period of the stimulus tone (T). Each histogram represents data accumulated over a thirty second interval. The vertical scales are plotted in units of spikes/second. The spontaneous rate, r_{sp}, of each unit is indicated by an arrow on the vertical scale. Each column contains period histograms computed from one unit, the lowest level at the top of the column. The stimulus level is specified as the rms sound pressure at the tympanic membrane, expressed in dB SPL. The characteristics of each unit are

Unit No.	r_{sp} (spikes/sec)	CF (kHz)
DJ35-41	74.5	0.21
DJ35-24	69.9	1.85
DJ35-38	77.2	3.70


UNIT
DJ 35-41
DJ35-24
DJ35-38
FREQUENCY OF TONE
0.213 kHz
1.815 kHz
3.409 kHz
RATE (SPIKES/SEC)
750
500
250
0
300
200
100
r_{sp}
LEVEL (dB SPL)
24.5
-6.0
8.5
34.5
4.0
18.5
44.5
14.0
28.5
54.5
24.0
38.5
64.5
34.0
48.5
74.5
44.0
58.5
0
T
T = PERIOD OF TONE

had similar spontaneous discharge rates and, for each fiber, the stimulus frequency was chosen to be near the CF. The relationship of the stimulus frequency to the CF of the fiber does not alter the waveform of the period histogram. Rather, the waveform of the period histogram is dependent upon the absolute stimulus frequency; the data illustrated in Figure 4.1 are representative of waveforms obtained in different frequency ranges.

In the left-hand column, period histograms of the responses to a 0.213 kHz tone are presented. At low levels* (24.5 dB and 34.5 dB), the waveforms resemble sinusoids, being symmetric about the spontaneous discharge rate. As the level is increased further (to the 44.5 dB level), the waveform becomes asymmetric: the peak of the waveform is narrower than the valley. In the valley, the rate of discharge is nearly zero. No significant skew is seen in this waveform. For higher levels (54.5 dB, 64.5 dB, and 74.5 dB), the waveform becomes skewed, the leading edge of the waveform having a much steeper slope than the trailing edge. The maximum rate attained by the rate of discharge is between 500 and 750 spikes/sec. The rate of discharge is also zero over a significant portion of the period. With the exception of "peak-splitting responses" (which will be described later), this sequence of

*In this section, the terms "low" and "high" levels indicate the amplitude of the stimulus relative to the level for which a detectable synchronized response can be obtained.

waveforms is typical of period histograms measured with stimulus frequencies below 1 kHz.

In the center column of Figure 4.1, the period histograms obtained from the responses of a fiber to a 1.815 kHz tone are found. For low to moderate levels (-6 dB, 4 dB, and 14 dB), these waveforms resemble those obtained with the 0.213 kHz tone (24.5 dB, 34.5 dB, and 44.5 dB) illustrated in the left column. These waveforms are symmetric about the spontaneous discharge rate at low levels with a tendency to become asymmetric at moderate levels. In contrast to the response pattern to the 0.213 kHz tone, the waveforms in the center column change relatively little as level is increased to higher levels (24 dB, 34 dB, and 44 dB). The waveforms show much less skew than the corresponding histograms in the left column. Note that the rate of discharge is not maintained at zero for any significant portion of the period. This sequence of waveforms is indicative of the responses obtained with stimulus frequencies between one and three kilohertz.

Finally, the response pattern to a 3.409 kHz tone is presented in the right column. These waveforms do not appear to deviate significantly from a sinusoid. The rate of discharge does not approach zero during any portion of the period. These waveform characteristics are representative of the response patterns to tones having frequencies greater than three kilohertz.

In summary, the deviation of the rate of discharge from a sinusoidal appearance is more pronounced for frequencies less than

1 kHz. In addition, the asymmetry and the skew of the response waveform become more significant as level is increased. In order to quantify these response properties, measures of the synchronized response and the average-rate response are presented in the next section.

4.2 Measurements of the Synchronized Response and the Average-Rate Response

The synchronization index S_f (defined in section 3.6) is used to measure the time-varying component of the response at the frequency f. In this section, the frequency f denotes the frequency of the stimulus tone. The synchronization index characterizes the response waveform well only if the amplitudes of higher harmonics in the rate of discharge are much smaller than the fundamental. Consequently, we shall examine the spectral composition of the response to a single tone.

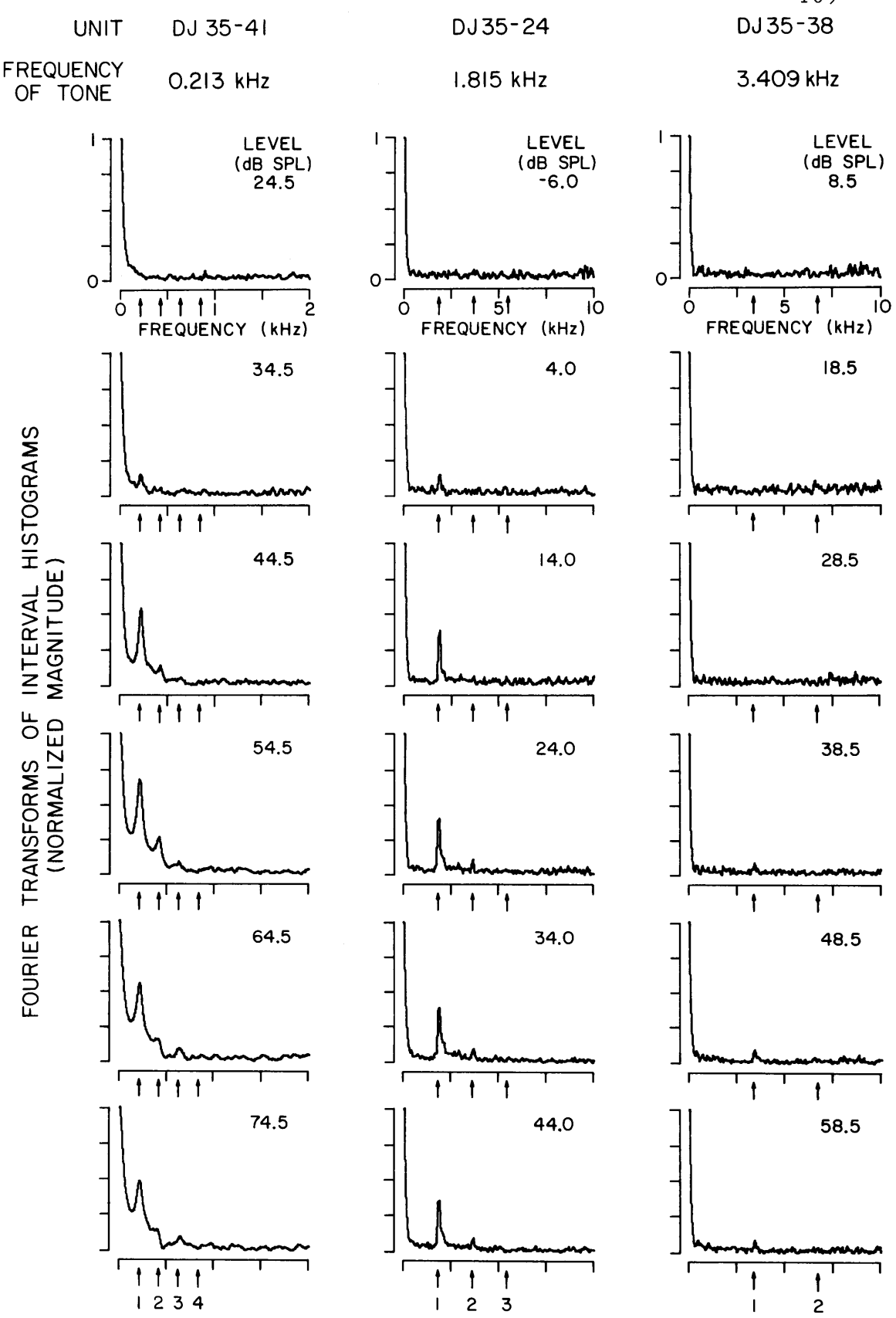
4.2.1 Fourier transforms of interval histograms computed from single-tone response patterns

The Fourier transform of an interval histogram displays the line spectrum of the rate of discharge (section 3.7). The amplitude of a component in this line spectrum is proportional to the square of the synchronization index measured at the frequency of the component. In Figure 4.2, the Fourier transforms of interval histograms computed from the responses to single tones are shown. The peaks present in these line spectra are those harmonically related to the stimulus. For each transform, the magnitude of the fundamental component (i.e., the first harmonic) is the largest component. The

Figure 4.2. Fourier Transforms of Interval Histograms Computed from the Responses to Single Tones

Each interval histogram was obtained from the same data used in computing the period histograms of Figure 4.1. The normalized magnitude of the Fourier transforms of these interval histograms are shown. The vertical scale is normalized so that the amplitude of the zero-frequency component is unity. The first, second, third, etc. harmonics of the stimulus frequency are indicated by an arrow on each frequency axis. The order of the harmonic is given in the bottom histogram of each column.

The interval histograms were computed with binwidths of 250 μsec for the first column and 50 μsec for the second and third columns. The computational details are given in section 3.8.


UNIT
DJ 35-41
DJ35-24
DJ35-38
FREQUENCY OF TONE
0.213 kHz
1.815 kHz
3.409 kHz
LEVEL (dB SPL)
24.5
-6.0
8.5
FREQUENCY (kHz)
34.5
4.0
18.5
44.5
14.0
28.5
54.5
24.0
38.5
64.5
34.0
48.5
74.5
44.0
58.5
FOURIER TRANSFORMS OF INTERVAL HISTOGRAMS (NORMALIZED MAGNITUDE)

presence of higher harmonics indicates that the rate of discharge is not a pure sinusoid. These higher harmonics are smaller in the 1.815 kHz response pattern than in the 0.213 kHz response pattern. No higher harmonics are seen in the 3.409 kHz responses. This progression toward fewer harmonics in the response pattern implies that there is less distortion of the rate of discharge as the stimulus frequency increases. This conclusion, derived from the spectrum of the response pattern, is consistent with the conclusions reached from examining the period histograms (Figure 4.1).

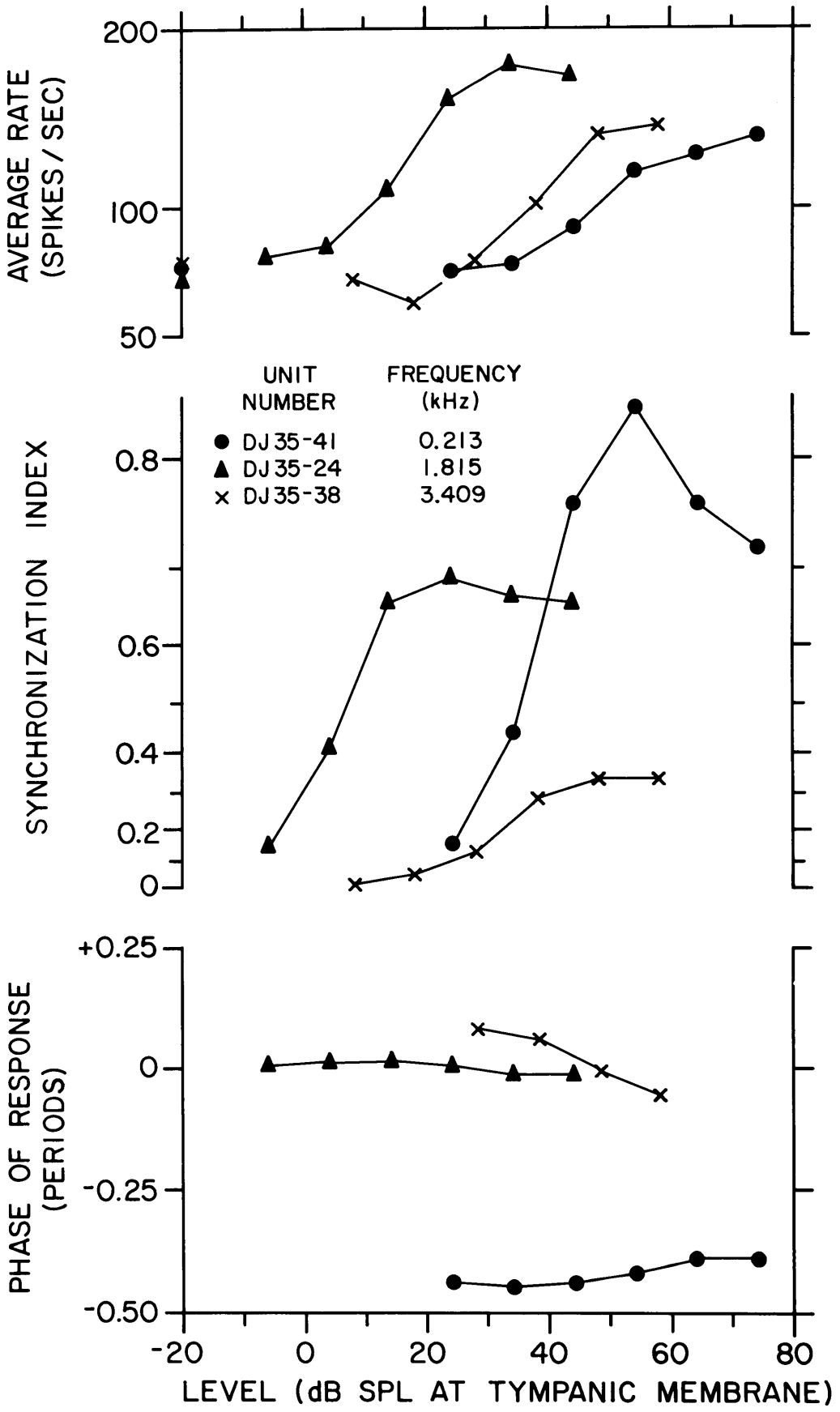
A complete specification of the response waveform would require a computation of the amplitudes of all the harmonic components. The amplitude of the fundamental (S_f) is, however, the largest of these components. The synchronization index reflects, therefore, the dominant component of the synchronized response pattern. However, one cannot expect to derive the waveform of the rate of discharge from S_f alone; the waveform depends upon the amplitudes of all of the harmonic components.

4.2.2 Characteristics of the fundamental component of the synchronized response

Response measures computed from the period histograms depicted in Figure 4.1 are shown in Figure 4.3. In this section, we discuss the synchronized response measures S_f and ϕ_f; the average-rate measurements are presented in the next section.

Figure 4.3. Measurements of the Average-Rate and the Synchronized Response from the Response Patterns to Single Tones

The measurements displayed in this figure were obtained from the period histograms depicted in Figure 4.1. The average rate, the synchronization index, and the phase of the response are displayed as a function of the rms sound pressure level at the tympanic membrane. The spontaneous discharge rates of the units are indicated on the left vertical axis of the upper panel. The synchronization index S_f and the phase of the response ϕ_f were computed according to the formulae given in equations 3.54a and 3.54b. Positive phase indicates a response leading the electrical output of the stimulus oscillator.


AVERAGE RATE (SPIKES / SEC)
200
100
50
UNIT NUMBER
FREQUENCY (kHz)
● DJ35-41 0.213
▲ DJ35-24 1.815
× DJ35-38 3.409
SYNCHRONIZATION INDEX
0.8
0.6
0.4
0.2
0
PHASE OF RESPONSE (PERIODS)
+0.25
0
-0.25
-0.50
-20
0
20
40
60
80
LEVEL (dB SPL AT TYMPANIC MEMBRANE)

For each of the three fibers we are considering, the synchronization index S_f rises monotonically with stimulus level until its maximum value is attained. The range of stimulus level required to increase S_f from a just-detectable value ($S_f > 0.1$) to a maximum value is 20-30 dB. The maximum value depends upon the frequency of the stimulus. For further increases of level, the synchronization index remains roughly constant for the 1.815 kHz and 3.409 kHz responses; however, S_f decreases significantly for the responses to the lowest frequency tone. This decrease in S_f with level is typical of the responses to single tones having a frequency less than 1 kHz.

Note that the values of S_f resulting from the 44.5 dB and 64.5 dB levels of the 0.213 kHz tone are virtually equal. However, the period histograms corresponding to these stimulus conditions are significantly different (see Figure 4.1). Apparently, the fundamental components of the response are the same in these two cases, but the harmonic content of the response differs (Figure 4.2).

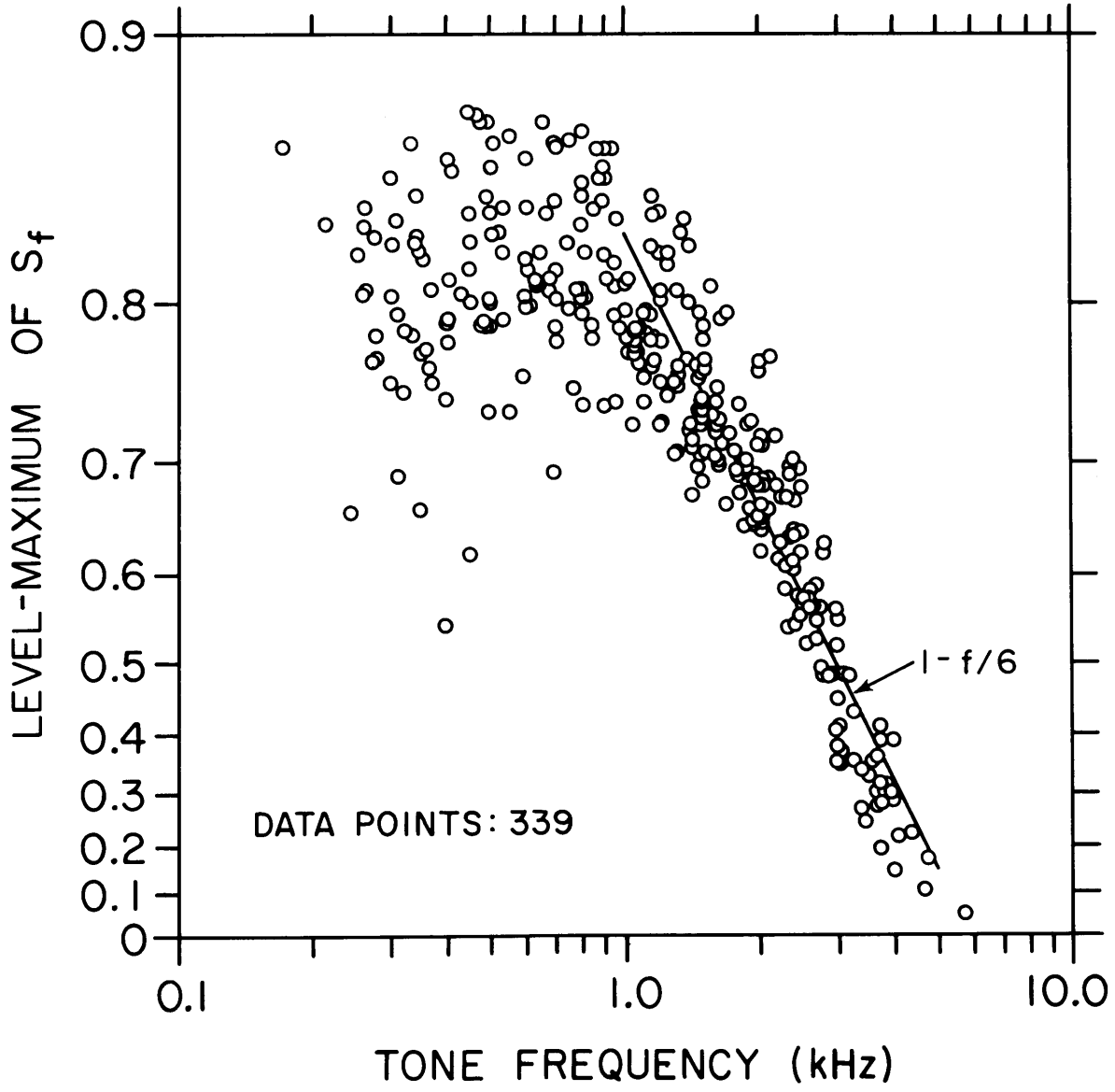
The maximum value attained by S_f as a function of level, termed the level-maximum of S_f, depends upon the stimulus frequency (Figure 4.4). In considering the variation of the level-maximum of S_f with stimulus frequency, the limits of synchronization possible in the response pattern are explored. This analysis was originally suggested by Dr. Julius L. Goldstein. Above 1 kHz, the level-maximum of S_f falls monotonically with stimulus frequency.

Figure 4.4. Level-Maxima of the Synchronization Indices Computed from the Responses to Single Tones

Level-maxima attained by synchronization indices measured from 339 level series taken from 233 units are displayed versus the frequency of the stimulus. The level-maximum of a response measure is a measurement of the largest possible value attainable by the response measure as a function of level, the level restricted to less than 80 dB SPL. For a particular level series, the level-maximum is defined to be equal to the largest value of the empirical response measures if a neighboring measurement (i.e., one taken within 10 dB in level) lay within plus or minus two standard deviations of the largest value. If this criterion was not satisfied, no level-maximum was defined. In the case of the synchronization index, an additional criterion was used; the largest measurement of S_f was accepted as the level-maximum if measurements of S_f for levels above and below were obtained. With this additional criterion, a level-maximum was defined for data exemplified by the measurements of S_f for the 0.213 kHz stimulus series of Figure 4.3.

The tone frequencies were not necessarily equal to the CF of the unit for the data depicted here. The line

(4.4) indicated on the figure represents the expression (1-f/6), f in kHz, where f is the tone frequency. This line is intended to represent the general trend of the data in this frequency range.


0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
LEVEL-MAXIMUM OF S_f
0.1
1.0
10.0
TONE FREQUENCY (kHz)
1 - f/6
DATA POINTS: 339

As the stimulus frequency nears 5 kHz, S_f becomes small (~0.1), making accurate measurement of S_f more difficult. A simple functional description of the level-maximum of S_f, max (S_f), in this frequency range is given by

$$\max (S_f) = 1 - f/6, \qquad 1 \leq f \leq 5 \tag{4.3}$$

where

f is the stimulus frequency in kHz.

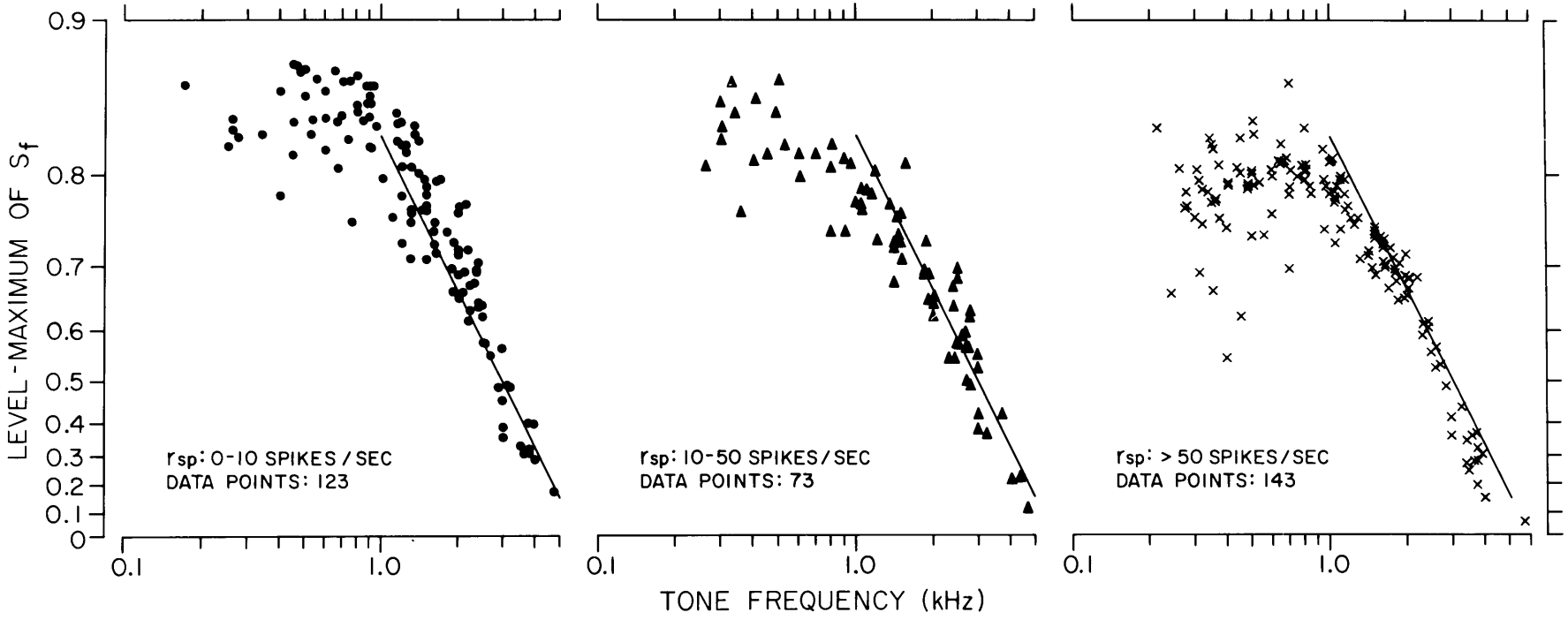
Below 1 kHz, the level-maximum of S_f is relatively constant. In this frequency range, the values of the level-maxima less than about 0.7 correspond to "peak-splitting" response patterns (see section 4.3 for a more complete discussion).

Below about 2 kHz, the spread in the data points displayed in Figure 4.4 is rather large. Some of this spread is apparently due to a small but systematic dependence of max (S_f) upon the spontaneous rates of discharge of the fibers (Figure 4.5). For frequencies below about 2 kHz, units with lower spontaneous rates tend to have larger values of the synchronization index.

The phase of the synchronized response, ϕ_f, is displayed in the bottom panel of Figure 4.3. This phase, measured for any one frequency, shows some variation with level. Such phase shifts can be more prominent if the stimulus frequency is not near the CF of the fiber. In our data, the amount of level-dependent phase

Figure 4.5. Level-Maxima of the Synchronization Indices Computed from the Responses to Single Tones: Variation with Spontaneous Rate

These data are replotted from Figure 4.4 on three axes based upon the spontaneous discharge rate (r_{sp}) of the fibers. The line drawn over the 1-5 kHz frequency range is given by the expression (1-f/6), f in kHz.


LEVEL-MAXIMUM OF S_f
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.1
1.0
rsp: 0-10 SPIKES / SEC
DATA POINTS: 123
rsp: 10-50 SPIKES/SEC
DATA POINTS: 73
rsp: > 50 SPIKES/SEC
DATA POINTS: 143
TONE FREQUENCY (kHz)

shifts is generally consistent with the data presented by Anderson et al. (1971). The relationship of the phase ϕ_f to the frequency of the stimulus for one unit was not systematically studied.

4.2.3 The average-rate response and its relationship to the synchronized response

The average discharge rate in response to a tone increases monotonically over a 20-30 dB range of stimulus level (see the upper panel of Figure 4.3). At the lower stimulus levels, a significant synchronized response can be present while the average rate remains near the spontaneous rate. The waveforms corresponding to these cases resemble sinusoids, symmetric about the spontaneous discharge rate (Figure 4.1). In general, a synchronized response can be detected some 10-20 dB below the level which evokes a just-detectable average-rate response.

For moderate levels, both the synchronization index and the average rate increase with stimulus level. In this range, the waveform of the rate of discharge becomes asymmetric for the two low-frequency stimuli (Figure 4.1). At the higher frequencies, the response pattern does not demonstrate asymmetric response waveforms for these stimulus levels.

For stimulus levels evoking response which have values of S_f near their level-maxima, the average rate of response may show a continuing increase. For instance, consider the measures of

the 1.815 kHz response shown in Figure 4.3. For the 14 dB and 24 dB stimulus levels, the synchronization index achieves its level maximum. The average discharge rate continues to increase in this range, achieving a maximum rate at the 34 dB level.

Pursuing the relationship between the synchronization index S_f and the average rate $\bar{r}$ further, the normalized values of $\bar{r}$ are plotted against the normalized values of S_f in Figure 4.6. The response characteristics of $\bar{r}$ and S_f define three response regions when viewed in this manner. If the normalized synchronization index is less than 0.8, the normalized average rate does not deviate significantly from zero. In other words, the average rate $\bar{r}$ is approximately equal to the spontaneous rate r_{sp} until the synchronization index reaches approximately 80% of its level-maximum. In the second region, the normalized average rate increases as the normalized synchronization index increases. Note, however, that the normalized synchronization index approaches 1 as the normalized average rate reaches 0.5. In the third response region, the normalized synchronization index remains near 1 while the normalized average rate spans 0.5 to 1. Consequently, the synchronization index remains near its level-maximum while the average rate spans the second half of its range.

Not only is the synchronization index S_f relatively constant in the third response region, but the response waveform, when adjusted for the average discharge rate, can also be constant. The rates of discharge r(t) of the responses to a 1.815 kHz tone are shown

Figure 4.6. Normalized Average Discharge Rate versus Normalized Synchronization Index: Single-Tone Response Patterns

For each stimulus presentation in a level series, the average discharge rate $\bar{r}$ and the synchronization index S_f were measured from the resulting response pattern. The average rate was normalized according to the expression

$$\text{norm}\,[\bar{r}] = (\bar{r}-r_{sp})/(\bar{r}_{max}-r_{sp})$$

where

$\bar{r}_{max}$ denotes the level-maximum of $\bar{r}$ from the level series

r_{sp} denotes the spontaneous discharge rate of the unit.

The synchronization index was normalized according to the formula:

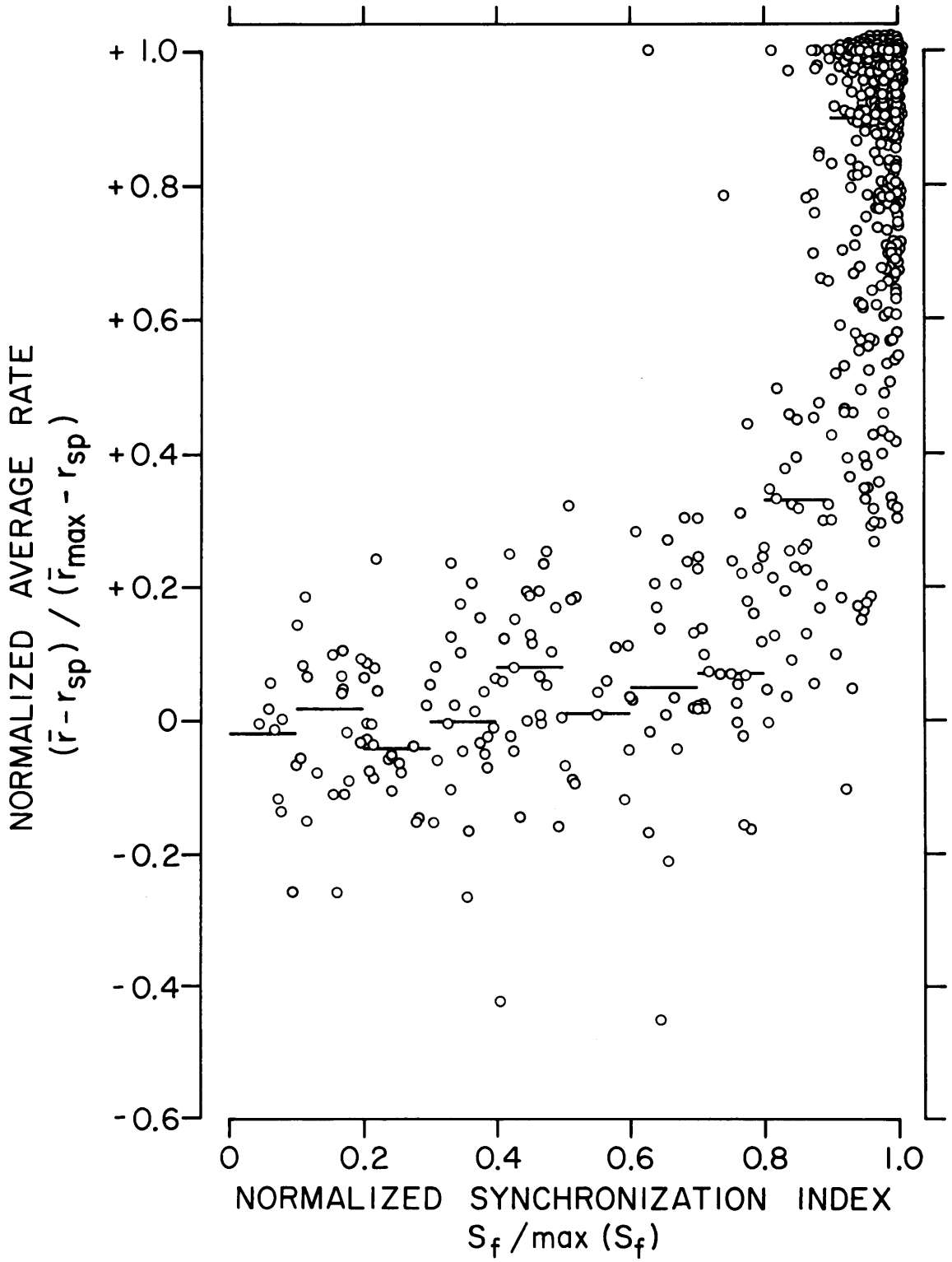
$$\text{norm}\,[S_f] = S_f/\max\,(S_f)$$

where

$\max\,(S_f)$ denotes the level-maximum of S_f from the level series.

These normalizations scale the response measures so that each spans the range 0 to 1 regardless of variation in spontaneous discharge rate r_{sp}, $\bar{r}_{max}$, and $\max\,(S_f)$.

The data obtained from a particular level series were included only if each response measure attained a level-maximum (84 level series). The lines indicate median values of norm $[\bar{r}]$ as measured over intervals of 0.1 of norm $[S_f]$.

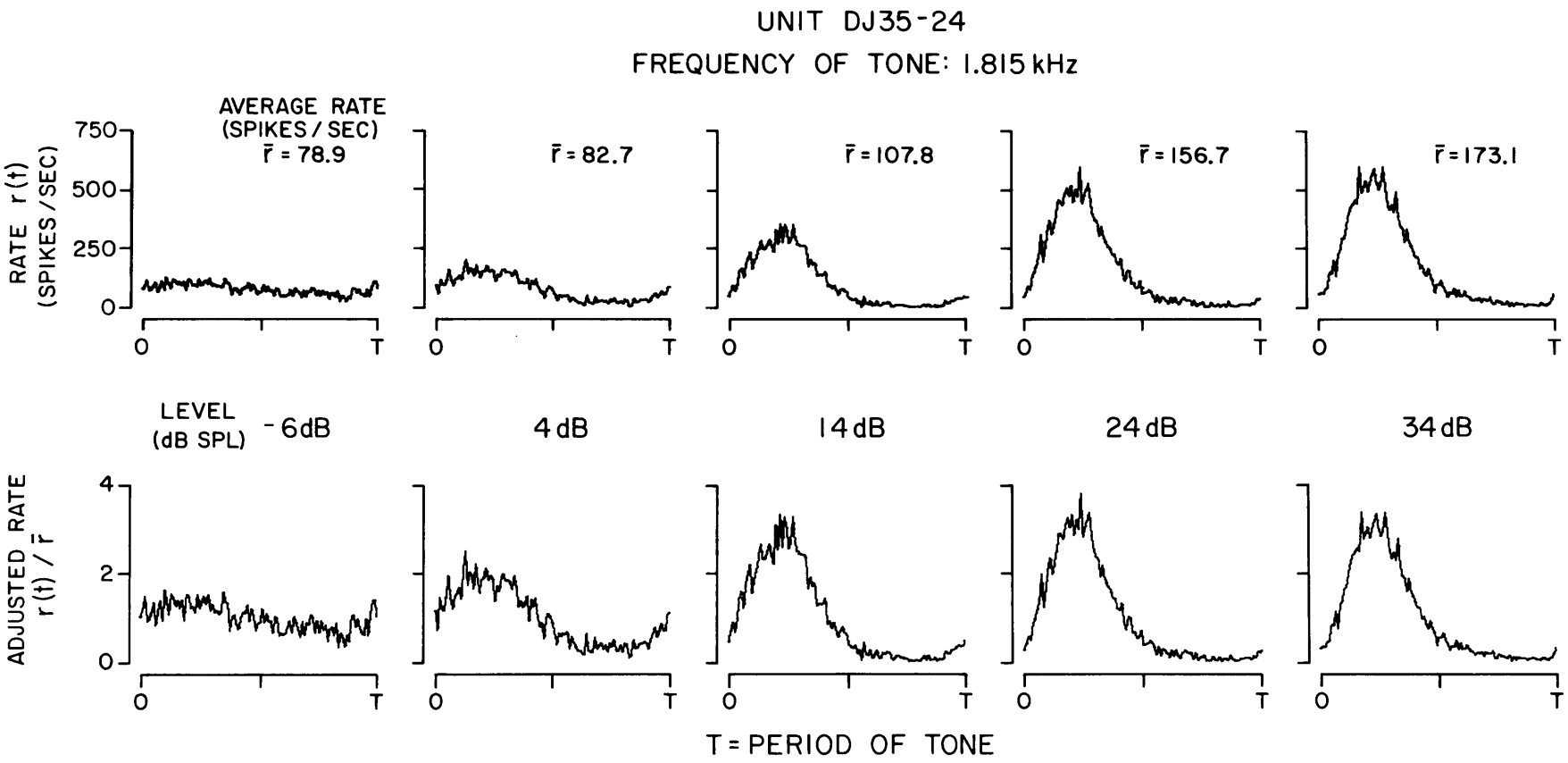

+ 1.0
+ 0.8
+ 0.6
+ 0.4
+ 0.2
0
- 0.2
- 0.4
- 0.6
NORMALIZED AVERAGE RATE
$(\bar{r} - r_{sp}) / (\bar{r}_{max} - r_{sp})$
0
0.2
0.4
0.6
0.8
1.0
NORMALIZED SYNCHRONIZATION INDEX
$S_f / max\ (S_f)$

in Figure 4.7 with the corresponding adjusted rates of discharge $r(t)/\bar{r}$. The adjusted rates obtained at the three largest levels resemble each other while the average discharge rates obtained at these levels differ significantly. Consequently, the response waveforms for the last three levels differ only by a multiplicative factor. This constancy of the single-tone response waveform in the third response region is typical for stimulus frequencies greater than 1 kHz.

In summary, the response measures S_f and $\bar{r}$ are monotonically related to stimulus level over approximately a 30 dB range. However, the overlap of these monotonic regions spans only a 10-20 dB range of levels. Consequently, some aspect of the response pattern to a single tone changes over a 40-50 dB range of levels providing the stimulus frequency is less than 5 kHz (so that a synchronized response can be measured).

Figure 4.7. Comparison of the Rate of Discharge of a Single-Tone Response Pattern with the Adjusted Rate of Discharge

The rate of discharge, r(t), and the adjusted rate of discharge, $r(t)/\bar{r}$, are shown for the response patterns recorded from one auditory-nerve fiber. The variable $\bar{r}$ is average discharge rate of the particular response. The top row contains period histograms; these period histograms are the same as shown in the middle column of Figure 4.1. The result of dividing the contents of each period histogram by the average discharge rate is shown in the bottom row. The stimulus level is specified as the sound pressure level at the tympanic membrane. The level increases from left to right in 10 dB increments. The horizontal scale of each histogram corresponds to one period (T) of the sinusoidal stimulus.


UNIT DJ35-24
FREQUENCY OF TONE: 1.815 kHz
AVERAGE RATE
(SPIKES / SEC)
r̄ = 78.9
r̄ = 82.7
r̄ = 107.8
r̄ = 156.7
r̄ = 173.1
RATE r(t)
(SPIKES/SEC)
750
500
250
0
LEVEL
(dB SPL)
-6dB
4dB
14dB
24dB
34dB
ADJUSTED RATE
r(t) / r̄
4
2
0
0
T
T = PERIOD OF TONE

4.3 The "Peak-Splitting" Response Pattern

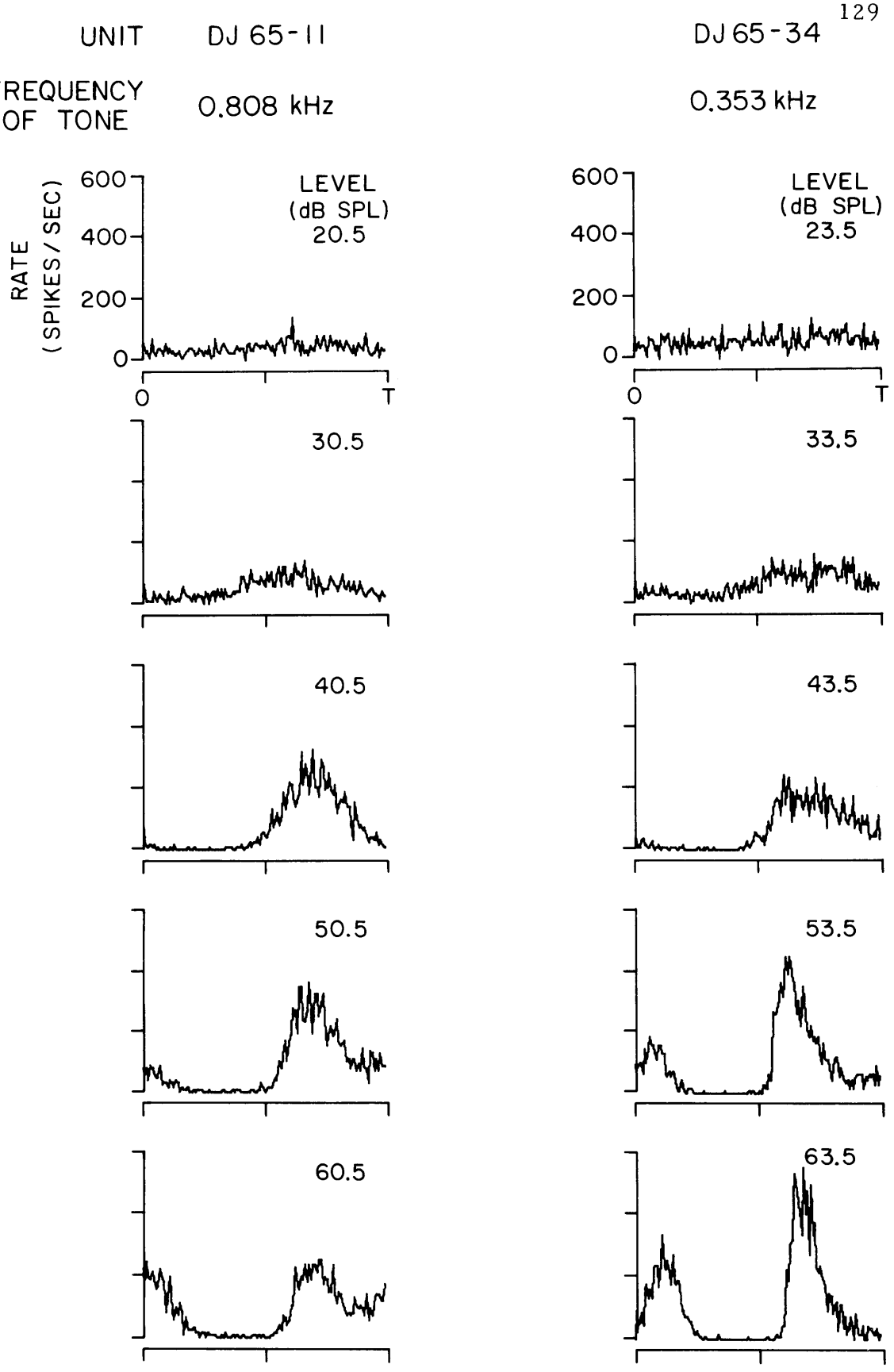
Period histograms computed from the response patterns to a single tone can have two dominant peaks per stimulus period. These "peak-splitting" response patterns have been studied by Dr. Nelson Y.-S. Kiang but his results have not been published. It is his idea that these patterns are a response to a second-harmonic of the stimulus tone. Examples of these "peak-splitting" response patterns are shown in Figure 4.8. For low to moderate levels, one peak per period is seen, the waveforms resembling those described in section 4.1. At higher levels, however, two peaks per period are obtained.

This pattern occurs only for stimulus frequencies less than 1 kHz (Kiang, 1974). Two distinct peaks per period are seen only if the stimulus level is greater than 50 dB SPL. This response pattern is not due to second harmonic distortion in the stimulus system. Measurements of the square-law distortion present in the acoustic output of the condenser earphones indicate that the amplitude of the second harmonic is -5.5 dB SPL and 3.5 dB SPL for the largest levels depicted in Figure 4.8 for the 0.808 kHz and 0.353 kHz responses respectively. The average rate thresholds of the fibers at the second harmonics of these frequencies are 21 dB SPL and 13 dB SPL respectively. Consequently, only a very small synchronized response to the second harmonic in the

Figure 4.8. Period Histograms Computed from "Peak-Splitting" Response Patterns

Period histograms computed from the response patterns of two auditory-nerve fibers are shown as a function of level. Each column contains the histograms from one unit, level increasing from the top to the bottom of each column. Stimulus level is specified as the sound pressure at the tympanic membrane. The vertical scale is expressed in spikes/second while the horizontal scale corresponds to one period of the stimulus (T). A "peak-splitting" response pattern is denoted by a period histogram having two prominant peaks per stimulus period. The characteristics of the units used in this figure are

Unit No.	r_{sp} (spikes/sec)	CF (kHz)
DJ65-11	38.7	1.52
DJ65-34	60.8	0.80


UNIT
DJ 65-11
DJ 65-34
FREQUENCY OF TONE
0.808 kHz
0.353 kHz
RATE (SPIKES/SEC)
600
400
200
0
LEVEL (dB SPL)
20.5
23.5
30.5
33.5
40.5
43.5
50.5
53.5
60.5
63.5
0
T
T = PERIOD OF TONE

acoustic stimulus would be expected.

The Fourier transforms of interval histograms measured from "peak-splitting" response patterns are depicted in Figure 4.9. As level is increased the amplitude of the fundamental first increases then decreases once the level-maximum is reached. The fundamental component of the response can become smaller than the second harmonic (right column of Figure 4.9). Consequently, the synchronization index S_f clearly does not adequately describe the "peak-splitting" response pattern.

Figure 4.10 shows measurements of the average rate and the synchronization index obtained from these response patterns. The variation of the average discharge rate with level resembles those depicted in Figure 4.3. [Although the maximum average rates are somewhat smaller in these instances, this is not an attribute of "peak-splitting" response patterns. "Peak-splitting" response patterns can have larger average discharge rates than shown in Figure 4.10 and non-"peak-splitting" responses can have smaller values of the maximum average rate.] The synchronization index decreases rapidly with level above the level of the level-maximum. The level-maxima in these instances are smaller than those obtained from non-"peak-splitting" response patterns. This result is consistently found in measurements obtained from "peak-splitting" responses. The data points in Figures 4.4 and 4.5 lying below 0.7 for frequencies less than 1 kHz were obtained from "peak-splitting" response patterns.

Figure 4.9. Fourier Transforms of Interval Histograms Computed from "Peak-Splitting" Response Patterns

Interval histograms were computed from the same data from which the period histograms of Figure 4.8 were computed. Magnitudes of the Fourier transforms of there interval histograms were calculated and normalized so that the magnitude of the zero-frequency component was unity. The binwidth of the interval histograms was 200 μsec. Computational details are given in section 3.8.

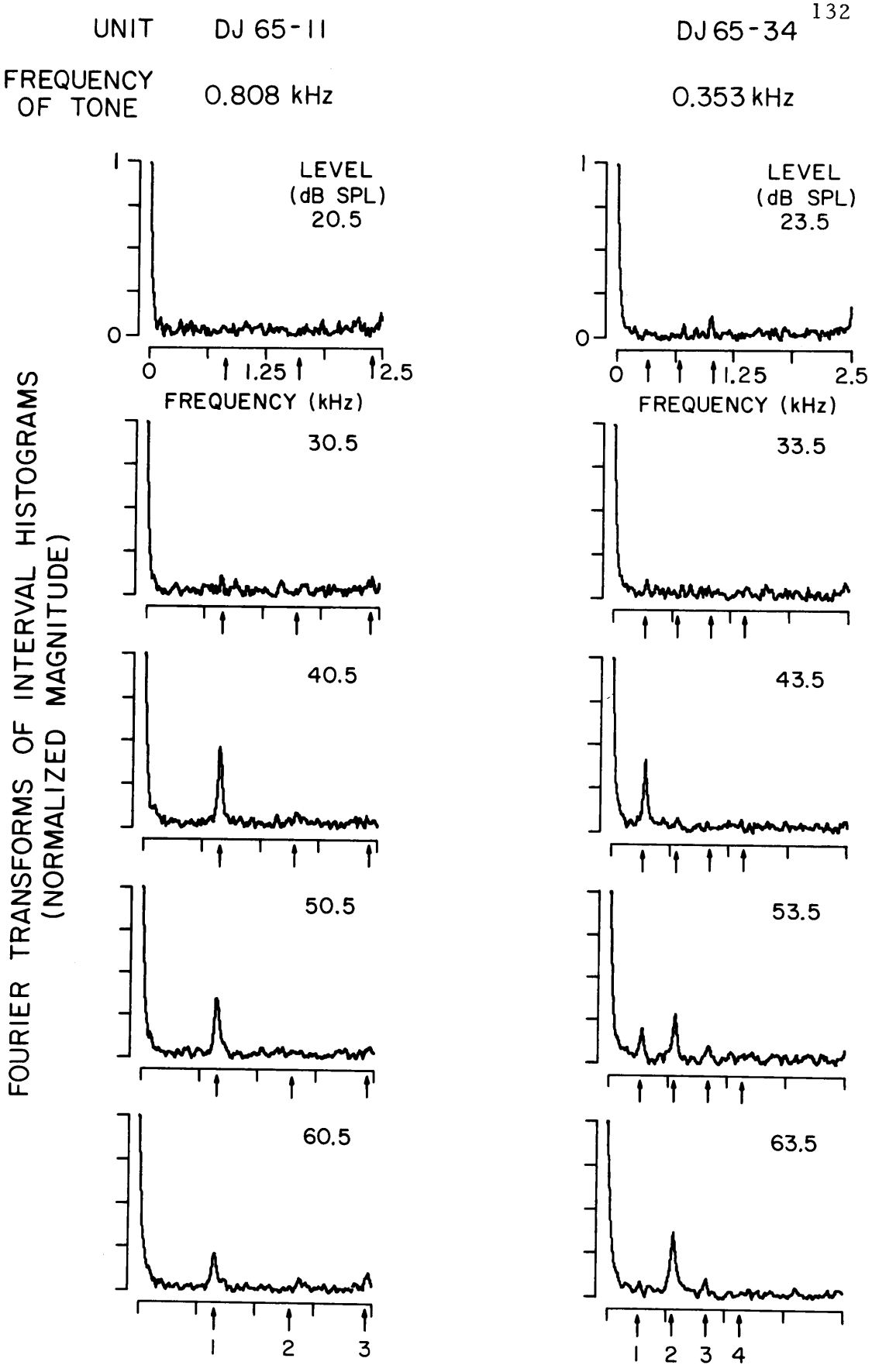
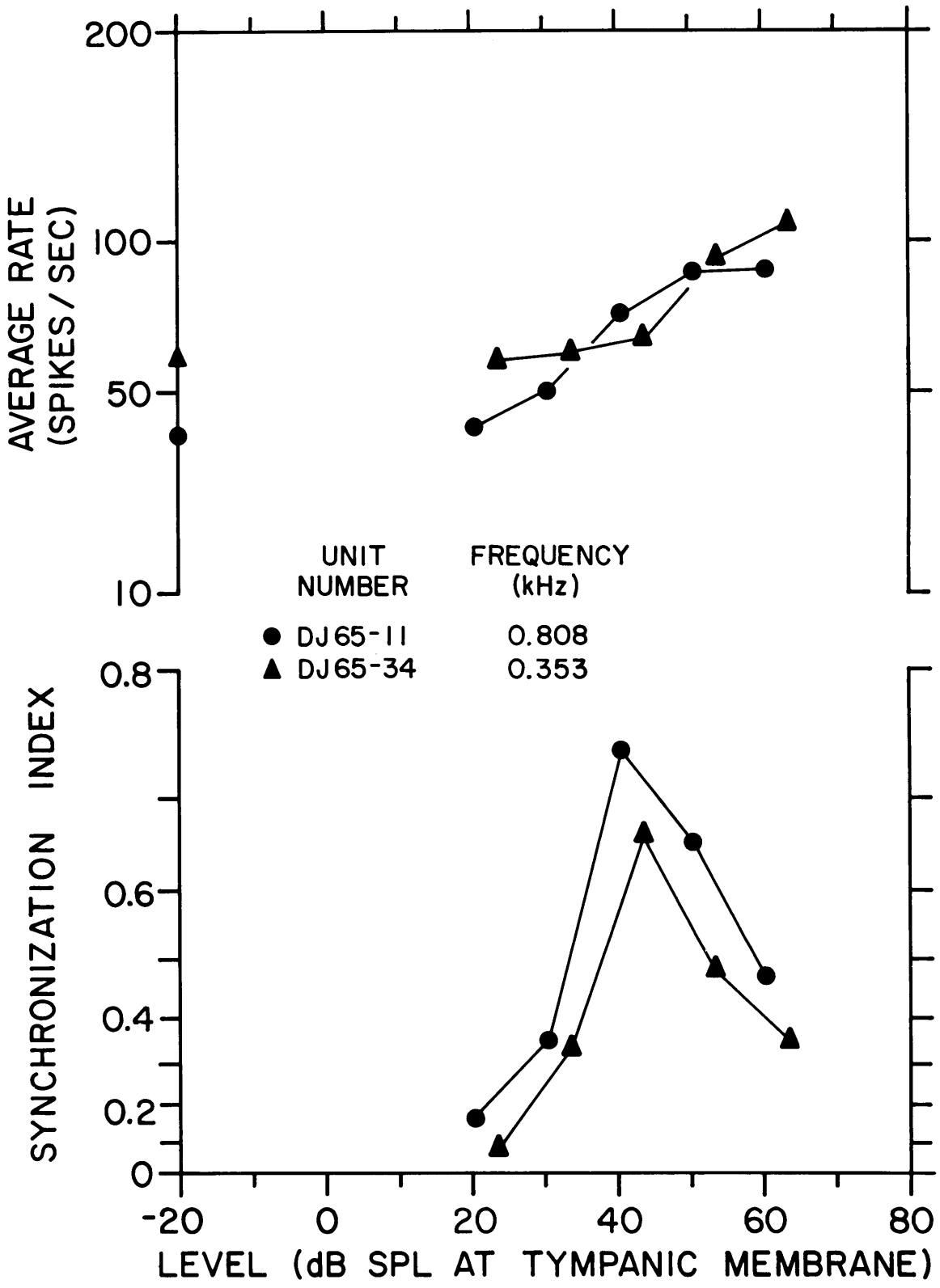

UNIT
DJ 65-11
DJ 65-34
FREQUENCY OF TONE
0.808 kHz
0.353 kHz
LEVEL (dB SPL) 20.5
LEVEL (dB SPL) 23.5
1
0
0
1.25
2.5
FREQUENCY (kHz)
30.5
33.5
40.5
43.5
50.5
53.5
60.5
63.5
1
2
3
4
FOURIER TRANSFORMS OF INTERVAL HISTOGRAMS (NORMALIZED MAGNITUDE)

Figure 4.10. Average Discharge Rate and Synchronization Index Measured from "Peak-Splitting" Response Patterns

The average discharge rate and the synchronization index are plotted as a function of stimulus level for the period histograms of Figure 4.8. The spontaneous discharge rate of each unit is indicated on the left vertical axis of the upper panel.


200
100
50
10
AVERAGE RATE
(SPIKES / SEC)
UNIT
NUMBER
FREQUENCY
(kHz)
● DJ65-11
0.808
▲ DJ65-34
0.353
0.8
0.6
0.4
0.2
0
SYNCHRONIZATION INDEX
-20
0
20
40
60
80
LEVEL (dB SPL AT TYMPANIC MEMBRANE)

4.4 An Analytical Description of the Rate of Discharge

In the previous sections of this chapter, the response patterns of auditory-nerve fibers to a single tone have been presented. The measurements of these responses provide the empirical basis of a model describing the response pattern to a tone. The model is derived by a two-step procedure.

1. The rate of discharge resulting from a single-tone stimulus is approximated by an analytical function. The parameters of this approximation are estimated directly from measurements of the rate of discharge. The variation of these parameters with stimulus level stimulus frequency is then determined.
2. A model is presented which describes the behavior of these parameters with variations in the stimulus.

This section is concerned with the first step in this procedure. Section 4.5 contains the presentation of the model.

4.4.1 Approximation of the rate of discharge

The approximation used here to describe the rate of discharge of an auditory-nerve fiber in response to a tone of frequency f is given by

$$r(t) = \exp\{a_o + a_1 \cos 2\pi(ft+\theta_1) + a_2 \cos 2\pi(2ft+\theta_2)\} \qquad (4.4)$$

where

$r(t)$ denotes the rate of discharge.

a_1, θ_1 are the amplitude and phase respectively of a component having the same frequency as the stimulus.

a_2, θ_2 are the amplitude and phase respectively of a component having twice the frequency of the stimulus (i.e., the second harmonic).

a_o is independent of time.

We shall refer to this approximation as the exponential approximation. This approximation contains five parameters, all of which are allowed to vary with stimulus parameters. The time-varying aspect of the rate of discharge is due to the components at the first and second harmonics of the stimulus frequency. If the amplitude of the second harmonic, a_2, is zero and the amplitude of the first harmonic, a_1, is small, then the exponential approximation yields a waveform having a sinusoidal appearance. As a_1 increases, the waveform becomes asymmetrical due to the influence of the exponential upon the first-harmonic component. With the appropriate relationship of the phase of the second harmonic, θ_2, to the phase of the first harmonic, θ_1, an increase in a_2 results in a skewed waveform. If a_2 becomes comparable to a_1, a large second-harmonic component would be present in $r(t)$; in this manner,

"peak-splitting" response patterns could be described. The average rate of discharge of the exponential approximation depends upon a_o, a_1, and a_2. The time-independent component, having an amplitude equal to a_o, can alter only the average rate of discharge. Consequently, the average discharge rate can be manipulated without modifying the waveform of the exponential approximation.

The choice of equation 4.4 as an approximation to measured rates of discharge is, of course, arbitrary. This approximation is an extension of the one proposed by Siebert (1970). The capability of the exponential approximation to describe accurately measured waveforms will now be assessed.

The five parameters of equation (4.4) were derived from period histograms using maximum-likelihood estimation techniques; the details of the parameter estimation procedure are given in Appendix III. The parameters were estimated independently from each period histogram. No constraints were placed on the parameters; for instance, parameter values obtained at other frequencies and levels did not influence the estimation procedure. The parameters of the exponential approximation were computed for 1046 period histograms obtained from the response patterns of 111 units in 11 cats.

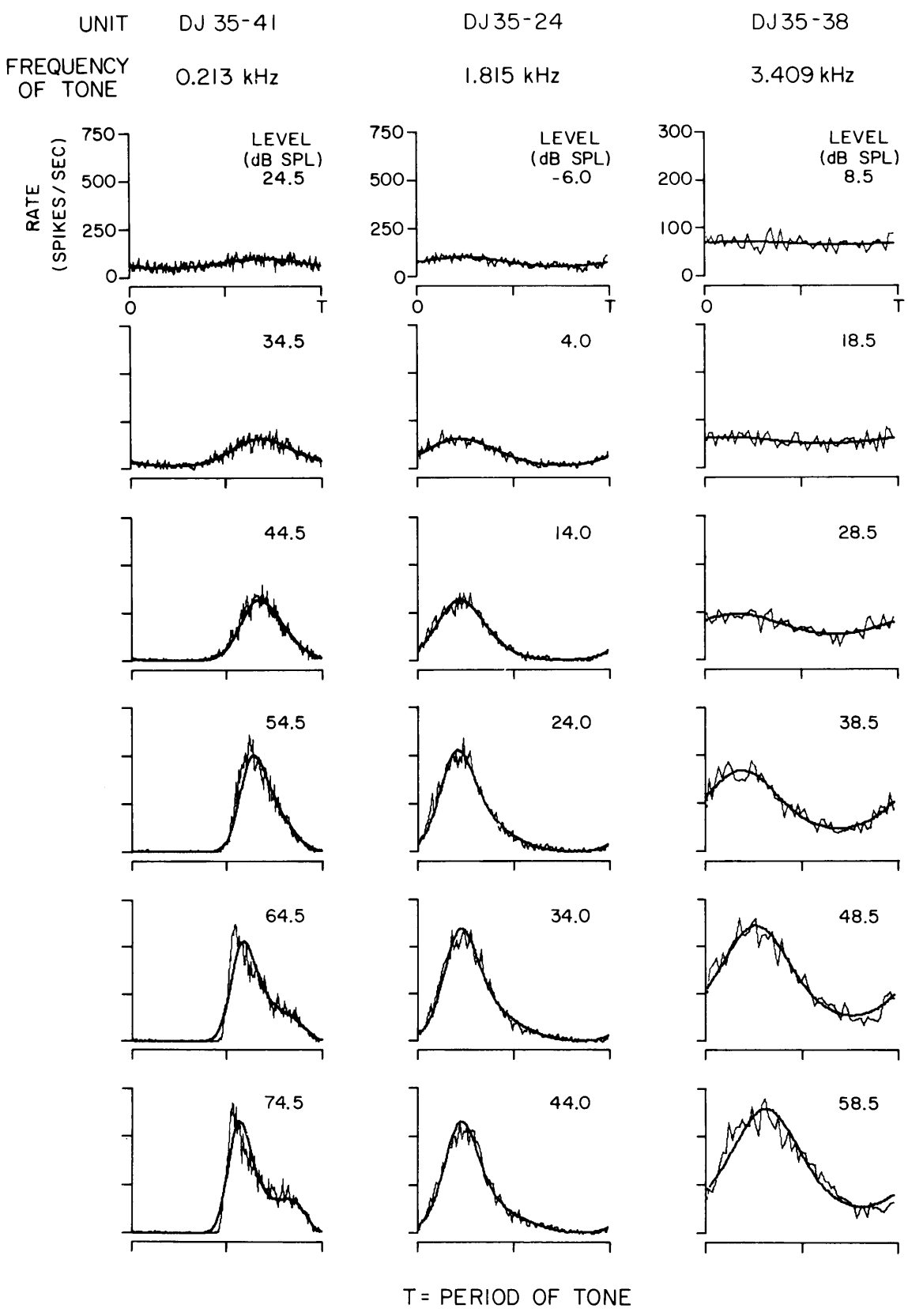
In Figure 4.11, examples of the results of this estimation procedure are depicted. The exponential approximations closely resemble the period histograms. The accuracy of the approximation, as portrayed in Figure 4.11, is representative of the results for all of the approximated period histograms. The largest deviation of the exponential approximation from a period histogram occur for high-level

Figure 4.11. Analytical Approximation of the Rate of Discharge for a Response to a Single Tone

The period histograms shown here are the same as those depicted in Figure 4.1. The analytical approximation to the rate of discharge is plotted on top of the period histograms in each case. The approximation used was

$$r(t) = \exp\{a_o + a_1 \cos 2\pi(ft+\theta_1) + a_2 \cos 2\pi(2ft+\theta_2)\}$$

The five parameters of this exponential approximation were computed separately for each histogram according to the maximum likelihood estimation procedure described in Appendix III.


UNIT
DJ 35-41
DJ35-24
DJ35-38
FREQUENCY OF TONE
0.213 kHz
1.815 kHz
3.409 kHz
RATE (SPIKES/SEC)
750
500
250
0
300
200
100
LEVEL (dB SPL)
24.5
-6.0
8.5
34.5
4.0
18.5
44.5
14.0
28.5
54.5
24.0
38.5
64.5
34.0
48.5
74.5
44.0
58.5
0
T
T= PERIOD OF TONE

and low-frequency stimuli (for example, the 64.5 dB and 74.5 dB levels of the 0.213 kHz stimulus series). In these instances, the mode of the approximation lags the mode of the period histogram and the leading edge of the approximation is not as steep as the data. The asymmetry and skew of the period histograms are well reflected in the approximations. No objective measure of "goodness-of-fit" was employed. The expression given by equation 4.4 appears to us to provide an adequate representation of the data.

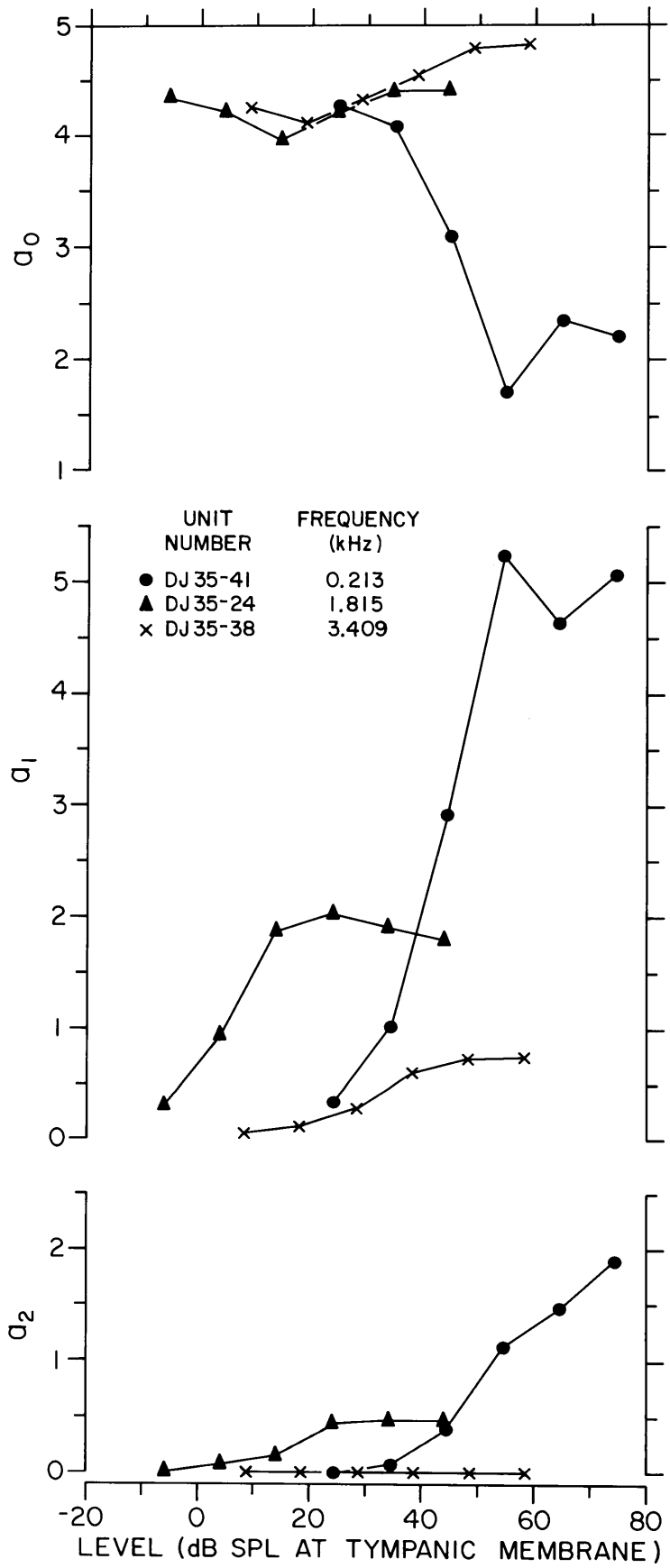
4.4.2 Variation of the Amplitude Parameters a_a and a_2 with stimulus parameters

The dependence of the model parameters a_o, a_1, and a_2 upon the level of the tone is illustrated in Figure 4.12. The amplitude of the fundamental-frequency component (a_1) increases with stimulus level over a range of approximately 20-30 dB until a maximum value is attained. For higher levels, a_1 maintains approximately the same value, this level-maximum depending upon the frequency of stimulus. This general description of the relationship between a_1 and stimulus level is valid for the frequencies exemplified in Figure 4.12 and for all of the analyzed data.

This dependence of a_1, the amplitude of the fundamental component in the exponential approximation, on stimulus level contrasts with measurements of the synchronization index, the amplitude of the fundamental in the response, shown in Figure 4.3. In the 0.213 kHz stimulus series, S_f decreases with increasing level once the level-maximum has been attained. This type of relationship is apparent only for stimulus frequencies less than 1 kHz. Referring to Figure 4.12, the values of a_1 corresponding to the decreasing values of S_f in the 0.213 kHz

Figure 4.12. Amplitude Parameters a_o, a_1, and a_2 of the Exponential Approximation to the Rate of Discharge: Variation with Stimulus Level

The values of the parameters a_o, a_1, and a_2 used to approximate the period histograms in Figure 4.11 are plotted versus the level of the stimulus tone. These parameters are dimensionless.


UNIT NUMBER
FREQUENCY (kHz)
DJ35-41 0.213
DJ35-24 1.815
DJ35-38 3.409
a_0
a_1
a_2
LEVEL (dB SPL AT TYMPANIC MEMBRANE)

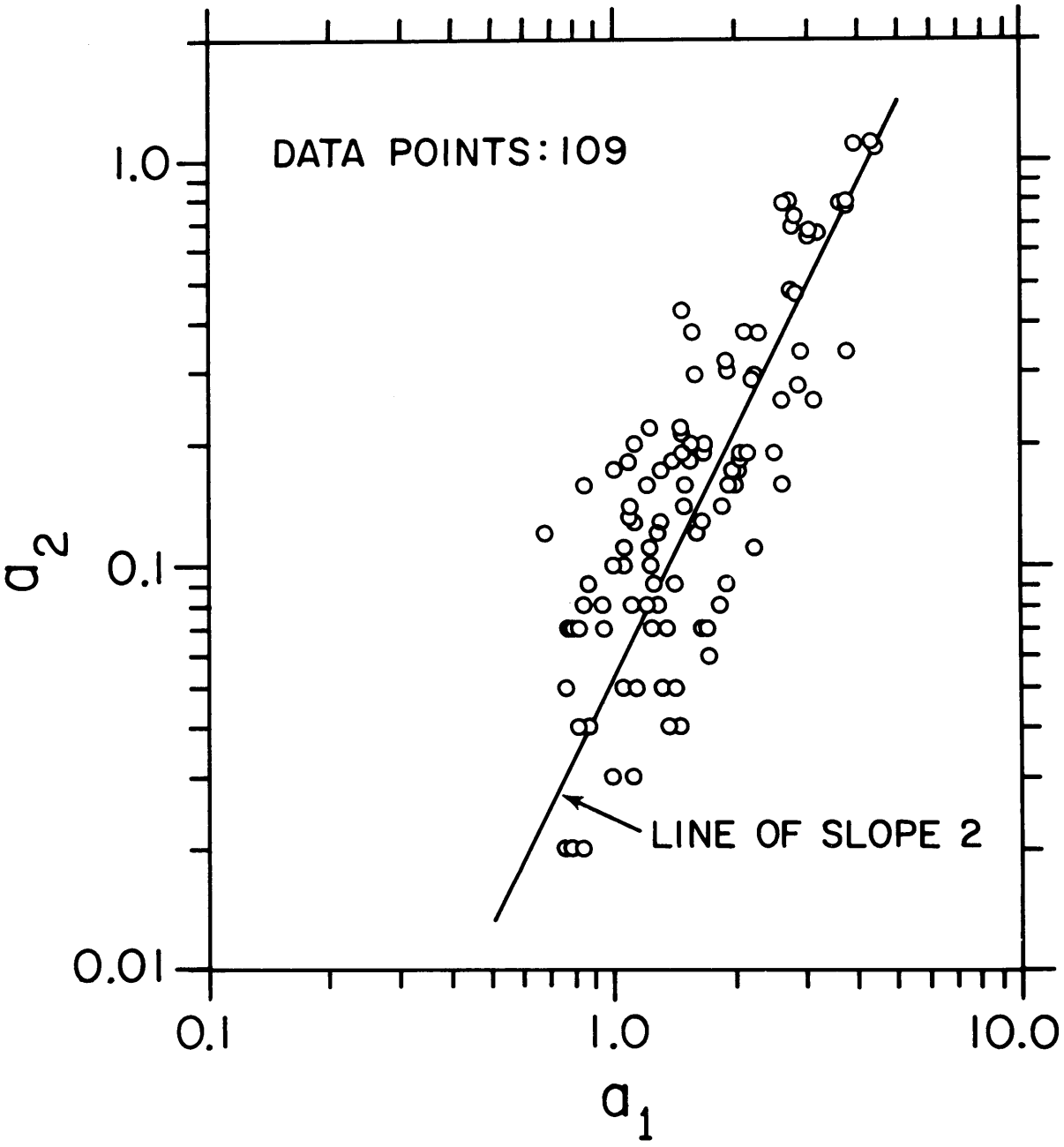
stimulus series are approximately equal (5.24, 4.62, and 5.07 for the 54.5 dB, 64.5 dB, and 74.5 dB stimulus levels respectively). These contrasting relationships of S_f and a_1 with stimulus level remains valid for stimulus frequencies less than 1 kHz.

In terms of the exponential approximations, S_f can decrease and a_1 remain constant if the amplitude of the second harmonic a_2 increases in this range. This behavior is exemplified in the 0.213 kHz stimulus series illustrated in Figure 4.12 (for the range of levels 54.5-74.5 dB). In this range of levels, S_f declines, a_1 remains approximately constant, and a_2 increases. Earlier, we related the decrease of S_f with increasing skewness of the waveforms of the period histograms. (See Figure 4.1.) Therefore, the increase of a_2 in the skew of the waveform and, because of the phase relationship between the first and second harmonics (section 4.4.3), a concomitant decrease in S_f. Measurable values of a_2 are also obtained in the 1.815 kHz stimulus series (Figure 4.12). However, these values of a_2 are small and the second harmonic component does not largely contribute to either the response waveform or the synchronization index for frequencies greater than 1 kHz.

The values of a_1 and a_2 appear to be related in a simple manner for the range of stimulus levels where both parameters increase monotonically. In Figure 4.13, values of a_2 are plotted against corresponding values of a_1. There is a considerable spread in the data points; however, a line of slope 2 on log-log coordinates describes the main trend in the data. Such a relation implies that a_2 is proportional to the square of a_1 in the range of levels for which each parameter is below its

Figure 4.13. Relation of the Amplitude Parameter a_2 to the Amplitude Parameter a_1

The values of the parameters a_1 and a_2 used in approximating period histograms measured from single-tone response patterns are plotted if the value of each parameter was less than its level-maximum by a factor of $1/\sqrt{2}$ (3 dB). No data points were included if a level-maximum for each parameter could not be defined.


DATA POINTS: 109
LINE OF SLOPE 2
1.0
0.1
0.01
a_2
0.1
1.0
10.0
a_1

level-maximum; i.e., $a_2 \propto a_1^2$.

The level-maxima attained by a_1 and a_2 depend upon the stimulus frequency (Figure 4.14). The raw data are shown in Figure 4.14a and third-octave averages of these data are given in Figure 4.14b. The level-maximum of a_1 is larger than the level-maximum of a_2 by a factor of 2-4, implying that the fundamental component dominates the second-harmonic component in the exponential approximations. If the level-maxima of a_1 and a_2 are plotted versus the frequencies of their respective components (Figure 4.14c), the dependence of these level-maxima upon frequency is approximately the same. They decline at a rate of 6 dB/octave between about 0.7 kHz and 2 kHz and fall at a faster rate above 2 kHz. The decline in the level-maximum of a_1 with stimulus frequency corresponds to the decrease of the level-maximum of S_f with increasing frequency.

4.4.3 Variation of the phase parameters θ_1 and θ_2 with stimulus parameters

The phase of the fundamental component, θ_1, in the exponential approximation is depicted in the upper panel of Figure 4.15 for the three stimulus series under consideration. The variation of θ_1 with stimulus level resembles the variation of the phase of the response ϕ_f. [The phases θ_1 and θ_2 of the exponential approximation are referred to a cosine (equation 4.4) while the phase of the response ϕ_f is referred to a sine (equation 3.47). If the phases of the fundamental components of the exponential approximation and of the response were equal, then $\theta_1 = \phi_f - 0.25$.] The phase angle $\theta_2 - 2\theta_1$ is shown in the

Figure 4.14. Level-Maxima of the Amplitude Parameters a_1 and a_2 Plotted versus Stimulus Frequency

The level-maxima of the parameters a_1 and a_2 of the exponential approximation are plotted on the same vertical scale versus stimulus frequency in panel A. The level-maximum of each parameter was defined to be equal to the maximum value of the parameter if either of two criteria were satisfied.

1. If the maximum measured value was obtained at the largest level used in the stimulus series, the value of the parameter obtained at the next lowest level must have been not less than 80% of the maximum value.
2. If the maximum value was not obtained at the largest level, the values obtained at higher levels could not be less than 80% of the maximum value.

If neither criterion was satisfied, no level maximum was defined. The level-maxima were defined independently for each parameter.

Panel B illustrates the average of the data in panel A computed in third-octave bands of frequency.

In panel C, the results shown in panel B are plotted versus the frequency of the component each parameter represents. In addition, the average values of the a_2 level-maxima were shifted upward by 7 dB, an amount which

(4.14) minimized the mean-square error between the level-maxima of a_1 and a_2. The solid line depicts an approximation to these data, which is given by

$$|H(f)| = \frac{6.5}{\left[1 + \left(\frac{f}{0.63}\right)^2\right]^{1/2} \cdot \left[1 + \left(\frac{f}{3}\right)^2\right]^{1/2}}$$

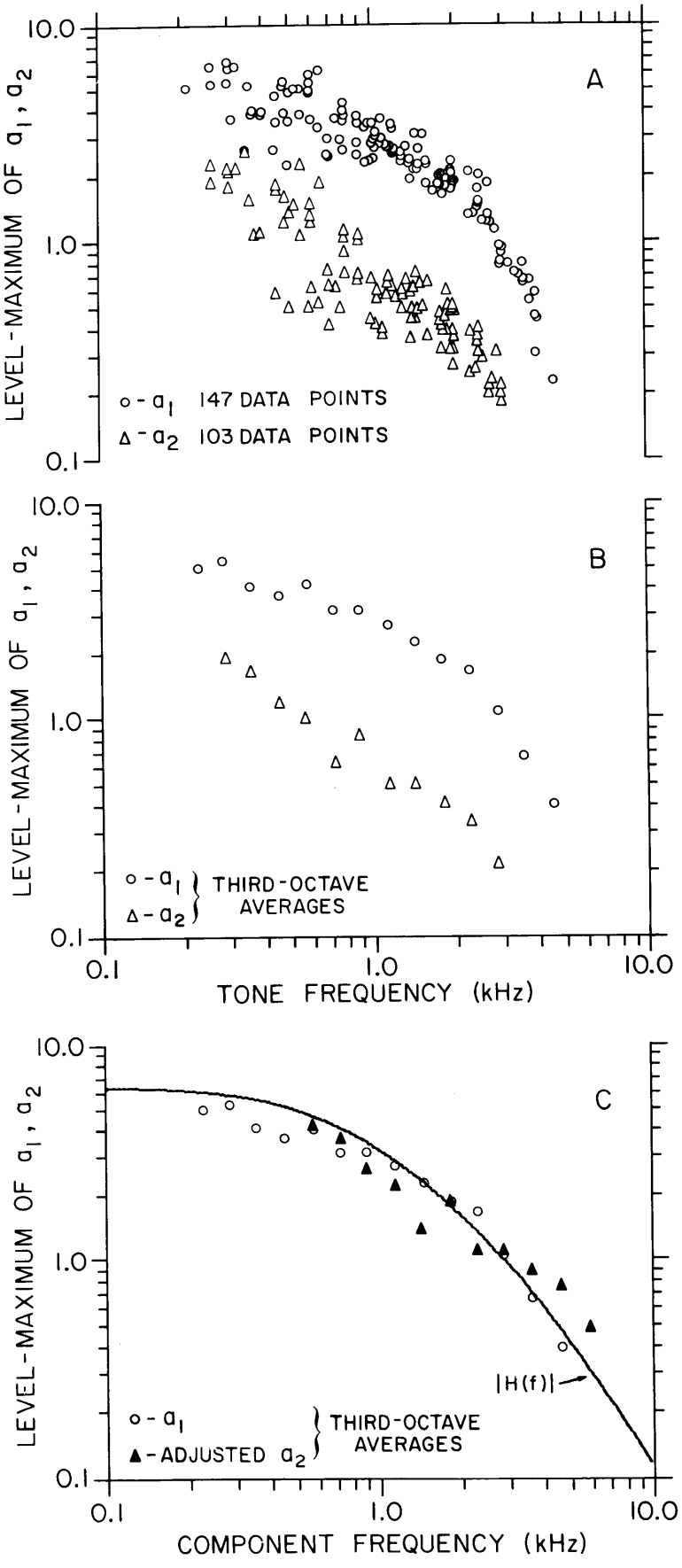
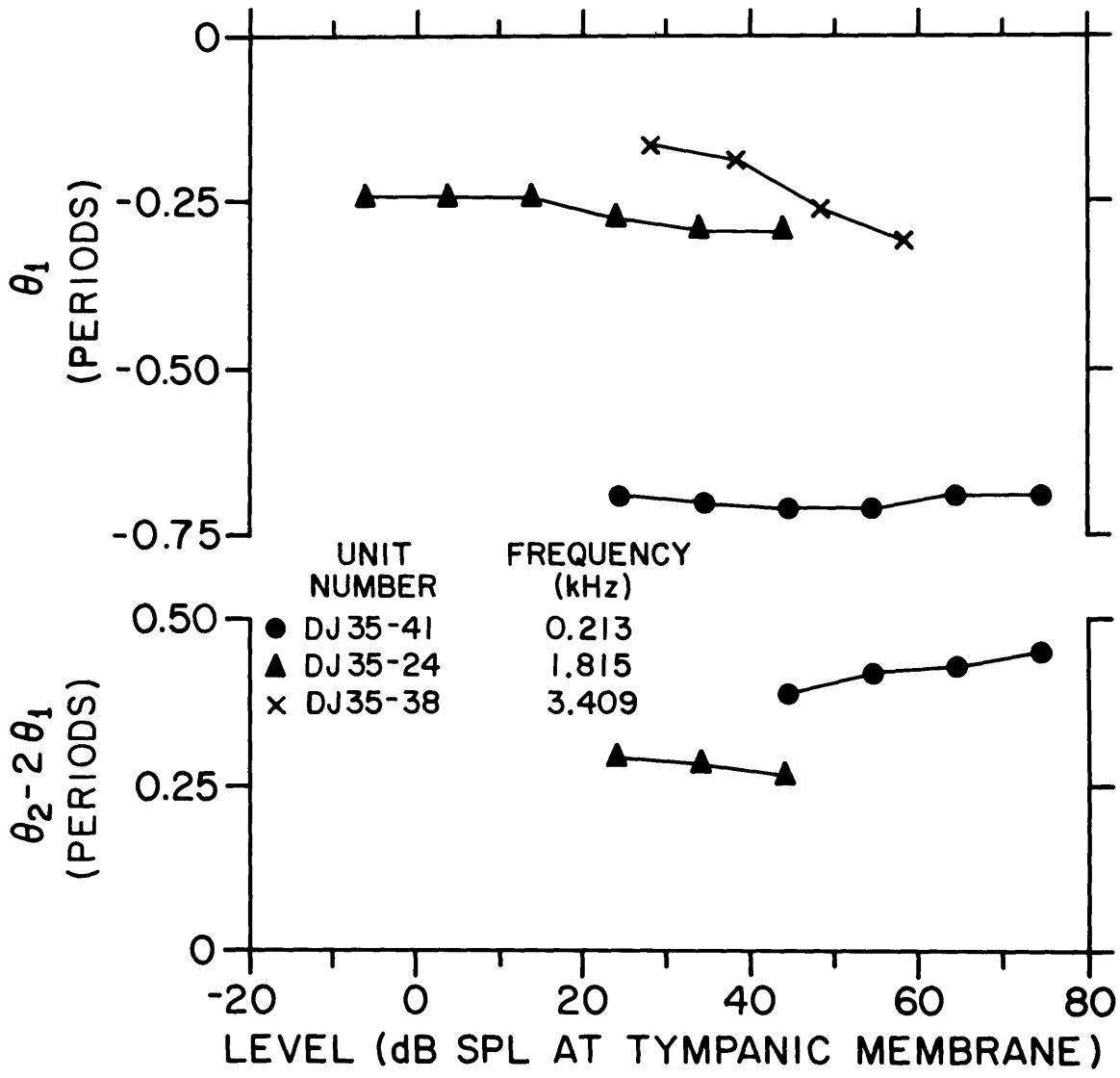

A
LEVEL-MAXIMUM OF a_1, a_2
10.0
1.0
0.1
o - a_1 147 DATA POINTS
Δ - a_2 103 DATA POINTS
B
LEVEL-MAXIMUM OF a_1, a_2
o - a_1 } THIRD-OCTAVE
Δ - a_2 } AVERAGES
0.1
1.0
10.0
TONE FREQUENCY (kHz)
C
LEVEL-MAXIMUM OF a_1, a_2
|H(f)|
o - a_1 } THIRD-OCTAVE
▲ - ADJUSTED a_2 } AVERAGES
0.1
1.0
10.0
COMPONENT FREQUENCY (kHz)

Figure 4.15. Variation of the Phase Parameters θ_1 and θ_2 with Stimulus Level

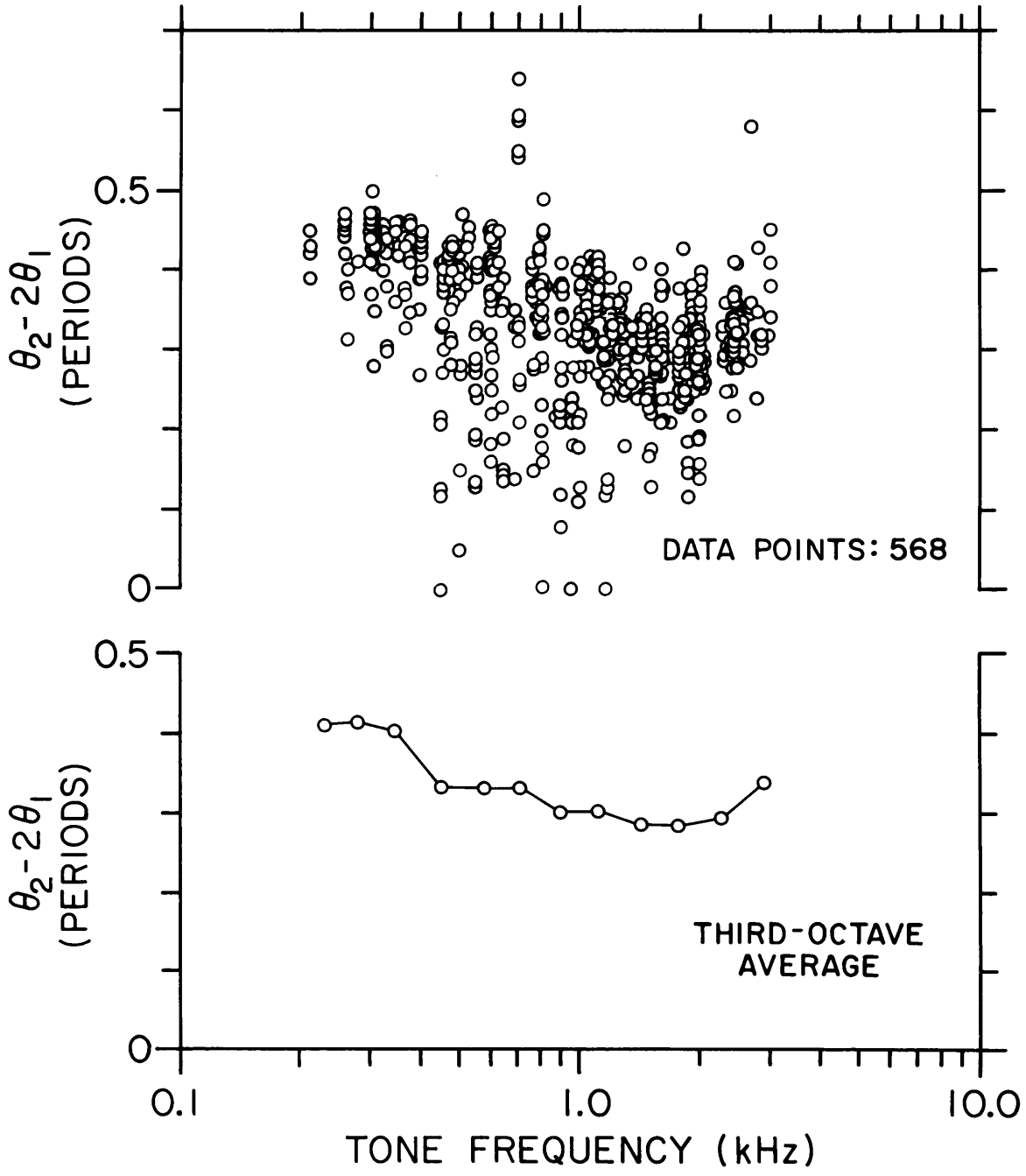
The phase angle of the fundamental component θ_1 in the exponential approximation is shown in the upper panel. The phase $\theta_2 - 2\theta_1$, the phase of the second-harmonic component relative to the fundamental, is displayed in the lower panel. These phase angles correspond to the values used in the exponential approximations depicted in Figure 4.11. Negative values of the phase angle θ_1 indicate lag with respect to the positive peak of the electrical output of the oscillator. Positive values of $\theta_2 - 2\theta_1$ indicate lead of the second harmonic with respect to the fundamental. Phase angles were plotted only if the corresponding amplitudes were greater than 0.2. As a result, $\theta_2 - 2\theta_1$, is not plotted for the highest-frequency data.


0
-0.25
-0.50
-0.75
θ_1 (PERIODS)
UNIT NUMBER
FREQUENCY (kHz)
● DJ35-41 0.213
▲ DJ35-24 1.815
× DJ35-38 3.409
0.50
0.25
0
$\theta_2 - 2\theta_1$ (PERIODS)
-20
0
20
40
60
80
LEVEL (dB SPL AT TYMPANIC MEMBRANE)

lower panel of Figure 4.15. This phase is the phase angle of the second harmonic relative to the fundamental. If $\theta_2-2\theta_1$ is zero, the second harmonic is in phase with the fundamental; if $\theta_2-2\theta_1$ is positive, the second harmonic leads the fundamental. For the examples shown in Figure 4.15, the second harmonic is nearly out-of-phase with the fundamental. The relative phase $\theta_2-2\theta_1$ can change with level: the data taken from the 0.213 kHz stimulus series demonstrates that with increasing level, $\theta_2-2\theta_1$ approaches 0.5. That is, the second harmonic approaches 180° out-of-phase with respect to the fundamental. Notice also that in the lower-frequency stimulus series, the second harmonic is more nearly 180° out-of-phase than in the 1.815 kHz series. This tendency appears to persist if we examine data from many units (Figure 4.16). The upper panel displays the data points of the phase angle $\theta_2-2\theta_1$ as a function of the stimulus frequency; the lower panel displays third-octave averages of these data. The spread in the raw data is great. This spread is due in part to level-dependent phase shifts. Despite the large spread, there is a tendency for the relative phase $\theta_2-2\theta_1$ to be nearer 0.5 at the lower frequencies ($f < 0.7$ kHz). In this frequency range, the largest shifts of $\theta_2-2\theta_1$ with level are obtained and the waveforms of the period histograms have the most skew. Therefore, with values of $\theta_2-2\theta_1$ slightly less than 0.5 with large values of a_2, the second harmonic component of the exponential approximation contributes skew to the rate of discharge, causing a sharper leading edge in the waveform than in the trailing edge.

Figure 4.16. Variation of the Relative Phase Angle $\theta_2 - 2\theta_1$ with Stimulus Frequency

The upper panel depicts measurements of $\theta_2 - 2\theta_1$ computed from 146 level series which were obtained from 94 units. No data point was included if the amplitude parameter a_2 was less than 0.2. The lower panel displays the average value of the data depicted in the upper panel. These averages were computed in third-octave frequency bands. Positive values of the phase $\theta_2 - 2\theta_1$ indicate lead of the second harmonic with respect to the fundamental.


0.5
0
$\theta_2 - 2\theta_1$ (PERIODS)
DATA POINTS: 568
0.5
0
$\theta_2 - 2\theta_1$ (PERIODS)
THIRD-OCTAVE AVERAGE
0.1
1.0
10.0
TONE FREQUENCY (kHz)

4.4.4 Variation of the amplitude parameter a_o with stimulus parameters

The component a_o in equation 4.4 affects only the average discharge rate; the average rate is also affected by the other four parameters, particularly a_1 and a_2. The parameter a_o is plotted versus stimulus level in the upper panel of Figure 4.12. When both a_1 and a_2 are near zero (e.g., low stimulus levels), the average rate of discharge (equal to the spontaneous rate r_{sp}) is determined solely by a_o. Consequently, for low levels a_o is given by

$$a_o = \ln r_{sp} \tag{4.5}$$

when

$$\bar{r} = r_{sp}, \quad a_1 = 0 = a_2.$$

As the level is increased, the value of a_o changes in a manner influenced by the frequency of the tone. For frequencies below about 0.5 kHz, a_o decreases with level until a minimum value is reached and then increases to somewhat larger values with further level increases (e.g., the 0.213 kHz stimulus series). The level at which the minimum value is reached corresponds to the level evoking the maximum value of S_f (Figure 4.3). For frequencies greater than 0.5 kHz, the relationship between a_o and stimulus level is roughly monotonic with the exception of statistical fluctuations. For the 1.815 kHz stimulus series in Figure 4.12, a_o remains relatively

constant with level, decreasing slightly for moderate levels. The parameter a_o increases to a final value larger than ln r_{sp} for the largest frequency stimulus. For frequencies greater than about 3 kHz, a_o is a monotonically increasing function of stimulus level.

While the value of a_o for low stimulus levels is determined by the spontaneous discharge rate, the final value reached by a_o as a function of level is relatively independent of the spontaneous rate of the fiber. Three stimulus series measured at approximately the same frequency are depicted in Figure 4.17 for units having different spontaneous rates. As level is increased, a_o approaches a value approximately the same for each of these fibers. Further evidence on this point is shown in Figure 4.18. The final value of a_o is plotted versus the frequency of the tone. These final values depend strongly upon the frequency of the stimulus. At any particular frequency, the final value of a_o is roughly independent of the spontaneous discharge rate; low spontaneous units tend to have slightly smaller final values of a_o than do high spontaneous units.

If the stimulus level is sufficiently large so that the level-maxima of a_1 and a_2 are attained as well as the final value of a_o, then, for stimulus frequencies above 1 kHz, the amplitude of a_o is larger than the amplitudes of either a_1 or a_2 (Figures 4.14a and 4.18). Above 1 kHz, the level-maximum of a_1 falls sharply with frequency; the final value of a_o, however, increases in this frequency range, approaching values near $a_o = 5$. As the time-varying components decline in amplitude, the time-independent

Figure 4.17. Dependence of the Amplitude Parameter a_o on Stimulus Level: Variation with Spontaneous Discharge Rate

Estimates of the amplitude parameter a_o are plotted against stimulus level for three different units. These data were chosen because similar stimulus frequencies were employed with units having different spontaneous rates. The characteristics of these units are:

Unit No.	r_{sp} (spikes/sec)	CF (kHz)	Tone Freq. (kHz)
DJ35-22	6.1	0.90	0.903
DJ35-8	44.3	1.00	1.002
DJ55-43	91.7	0.63	0.959

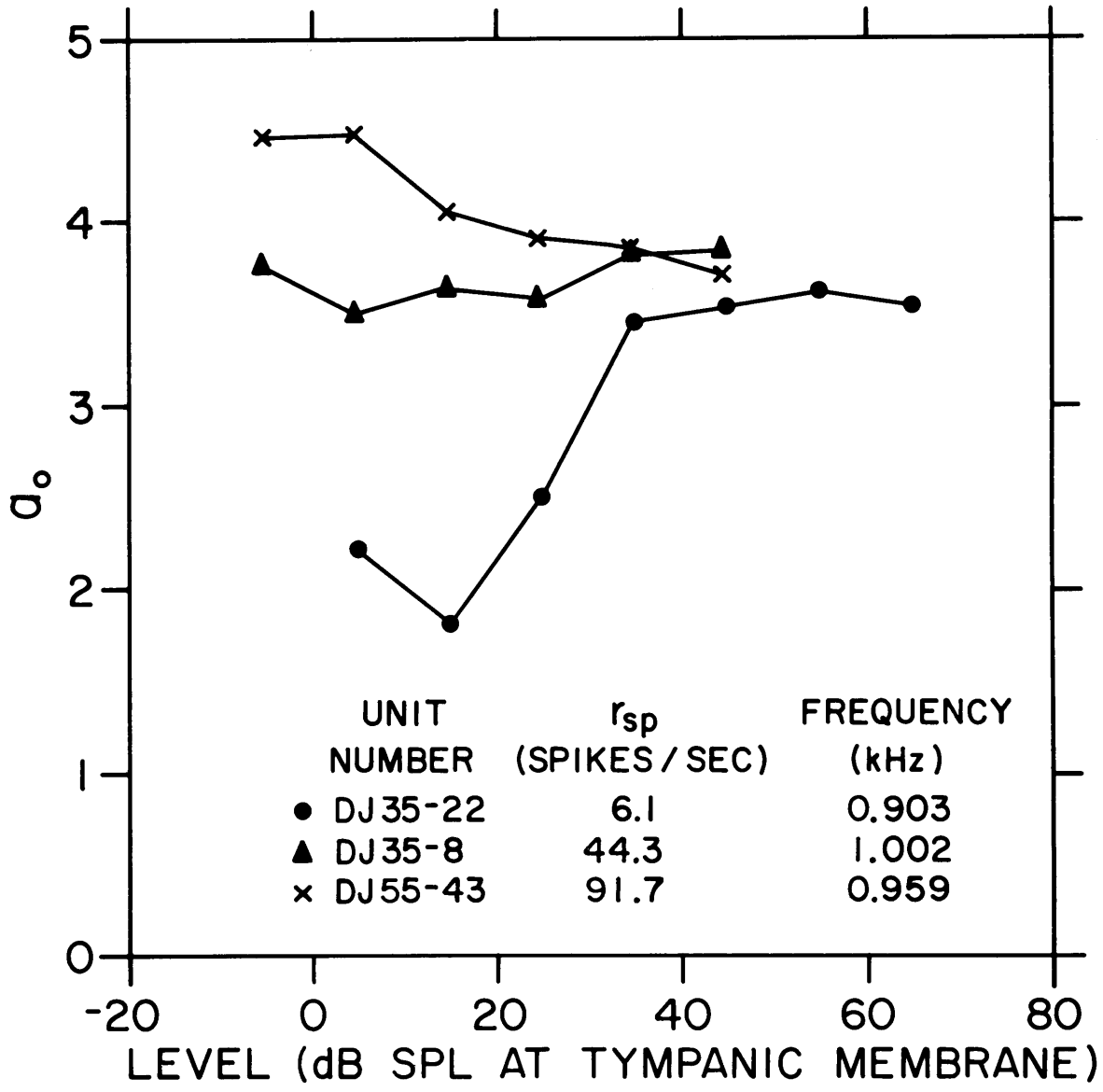
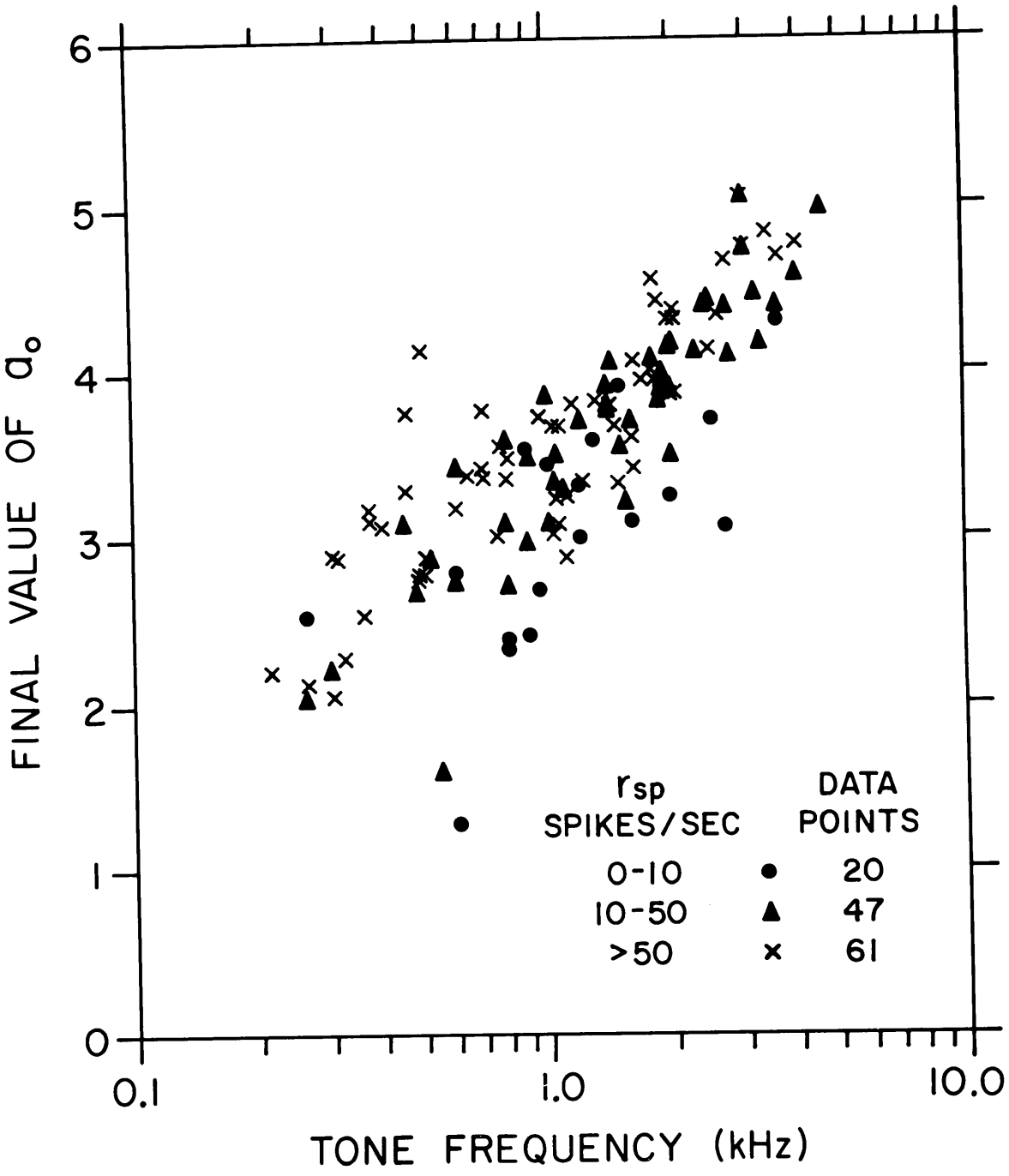

5
4
3
2
1
0
d°
-20
0
20
40
60
80
LEVEL (dB SPL AT TYMPANIC MEMBRANE)
UNIT NUMBER
r_{sp} (SPIKES / SEC)
FREQUENCY (kHz)
● DJ35-22 6.1 0.903
▲ DJ35-8 44.3 1.002
× DJ55-43 91.7 0.959

Figure 4.18. Final Value of the Amplitude Parameter a_o versus Stimulus Frequency: Variation with Spontaneous Discharge Rate

The relationship of the final value reached by the parameter a_o as a function of level is plotted against the stimulus frequency. The final value of a_o is the value of a_o measured at the largest stimulus level employed in a level series. Data points were included here only if the value of a_o at the next lowest level lay within 20% of the final value. The spontaneous discharge rates r_{sp} of the units are indicated by the different symbols.


FINAL VALUE OF a_o
TONE FREQUENCY (kHz)
0.1
1.0
10.0
0
1
2
3
4
5
6
r_{sp} SPIKES/SEC
DATA POINTS
0-10
20
10-50
47
>50
61

component a_o is the only significant term in the exponential approximation determining r(t). This is a reflection of the lack of a synchronized response at high frequencies (Figure 4.4) while an average-rate response remains.

4.4.5 Application of the exponential approximation to "peak-splitting" response patterns

Figure 4.19 displays a comparison of the "peak-splitting" response patterns depicted in Figure 4.8 with waveforms resulting from the exponential approximations. These results indicate that the period histograms can be fitted by the exponential approximations even when the rate of discharge exhibits "peak-splitting." The amplitude parameters a_1 and a_2 used in Figure 4.19 are plotted versus stimulus level in Figure 4.20. The variation of a_1 and a_2 with level is generally consistent with the trend mentioned earlier: these parameters increase with level, attain a maximum value, and remain approximately constant for higher levels. Note, however, that the level-maximum of a_1 is relatively small and the value of a_2 is somewhat large for these responses as compared with the entire body of data (Figure 4.14a). The amplitude of the second harmonic a_2 becomes comparable to a_1, the amplitude of the fundamental, resulting in two peaks per period in the period histograms.

A large decrease in S_f (Figure 4.10) occurs at levels for which the amplitude of the second harmonic increases while the amplitude

Figure 4.19. Analytical Approximation of the Rates of Discharge Obtained from "Peak-Splitting" Response Patterns

The exponential approximation expressed by

$$r(t) = \exp\{a_o + a_1 \cos 2\pi(ft+\theta_1) + a_2 \cos 2\pi(2ft+\theta_2)\}$$

is compared to period histograms computed from "peak-splitting" response patterns. These period histograms are identical to these illustrated in Figure 4.8. The parameters of the approximation were estimated independently for each histograms by the maximum likelihood estimation procedure outlined in Appendix III.

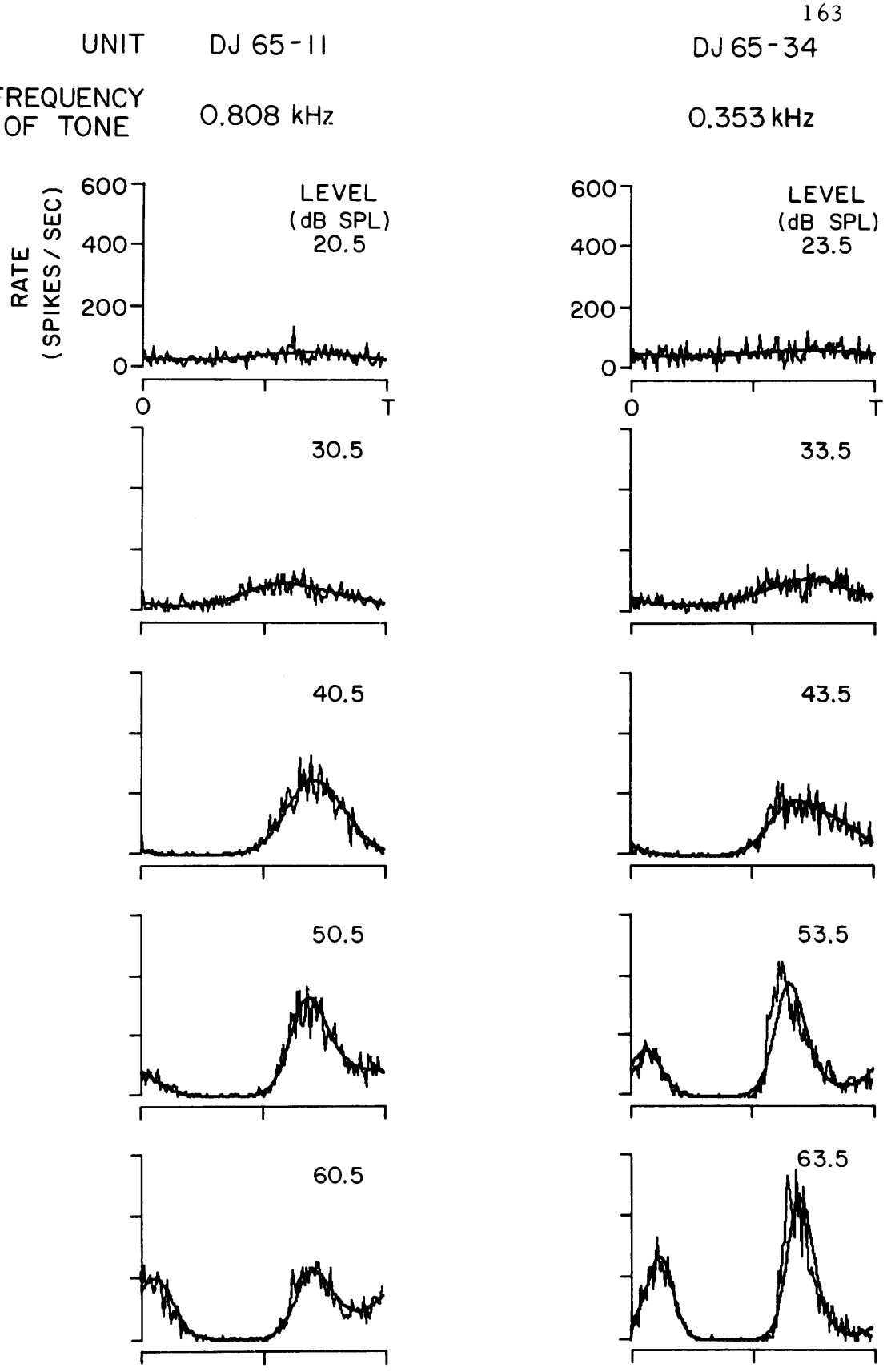
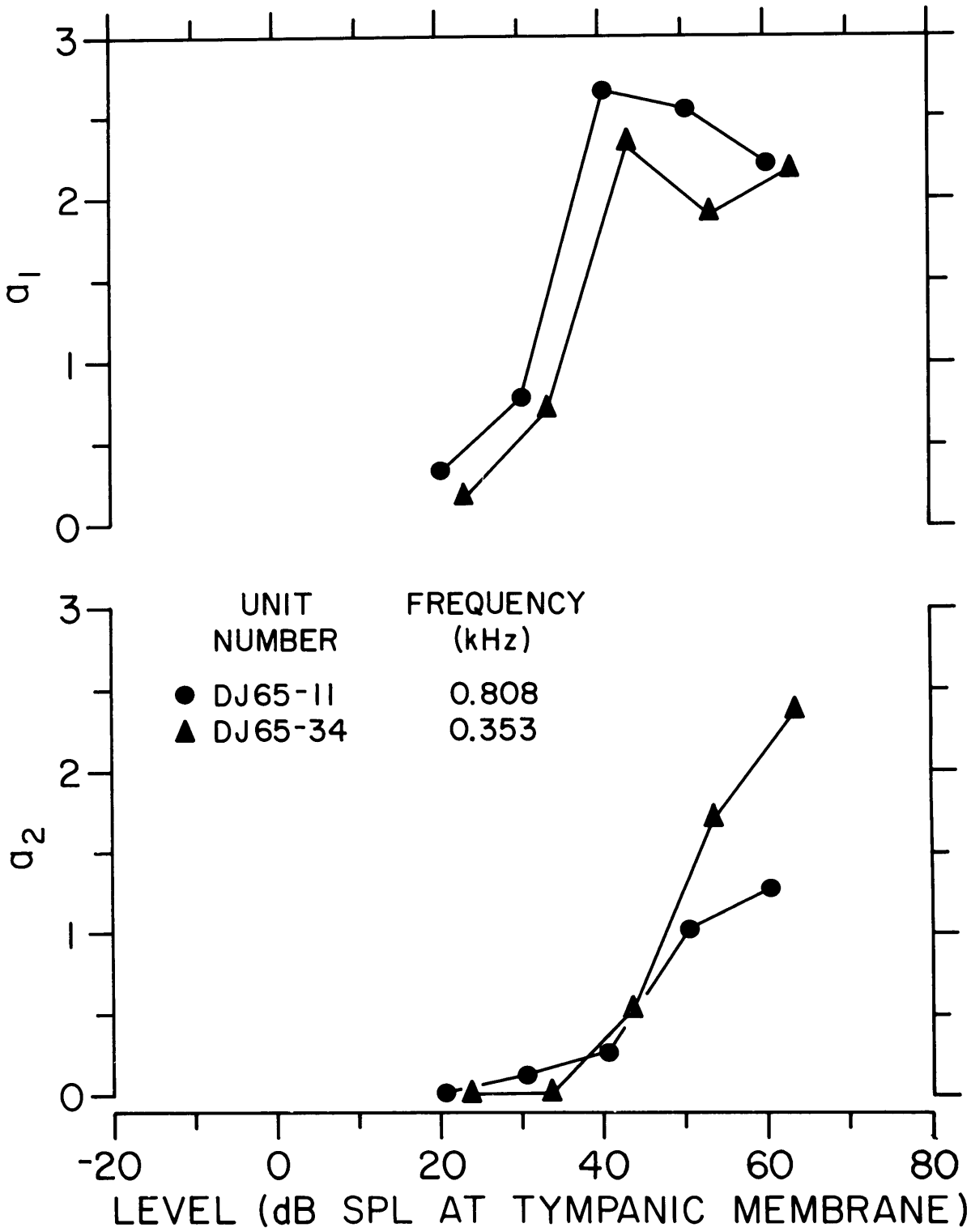

UNIT
DJ 65-11
DJ 65-34
FREQUENCY OF TONE
0.808 kHz
0.353 kHz
RATE (SPIKES / SEC)
600
400
200
0
LEVEL (dB SPL)
20.5
23.5
30.5
33.5
40.5
43.5
50.5
53.5
60.5
63.5
0
T
T= PERIOD OF TONE

Figure 4.20. Variation of the Amplitude Parameters a_1 and a_2 Obtained from "Peak-Splitting" Response Patterns with Stimulus Level

The values of the parameters a_1 and a_2 used in the exponential approximations depicted in Figure 4.19 are displayed as a function of stimulus level.


3
2
1
0
a_1
UNIT NUMBER
FREQUENCY (kHz)
● DJ65-11 0.808
▲ DJ65-34 0.353
a_2
-20
0
20
40
60
80
LEVEL (dB SPL AT TYMPANIC MEMBRANE)

of the fundamental remains relatively constant (Figure 4.20). This behavior can therefore be considered as an extension of the general trends described for non-"peak-splitting" response patterns (section 4.4.2). The difference between these response patterns is the relative amplitudes of a_1 and a_2 in the exponential approximation.

4.5 A Model of Single-Tone Response Patterns

The data show that the expression

$$r(t) = \exp\{a_o + a_1 \cos 2\pi(ft+\theta_1) + a_2 \cos 2\pi(2ft+\theta_2)\} \qquad (4.4)$$

accurately describes the rate of an auditory-nerve fiber responding to a tone of frequency f. The variation of the five parameters of this exponential approximation with stimulus parameters has also been described. In this section, we present an analytical model which describes the responses to single tones. Since equation 4.4 accurately fits the waveform of the response, the model need only describe the relation of the parameters of the exponential approximation to stimulus parameters. The model is intended to describe these relationships for a typical auditory-nerve fiber.

Level-dependent phase shifts of the fundamental component of the exponential approximation, θ_1, are not described by this model. This effect was not systematically studied in the course of this research; therefore, the model describes those response aspects measured directly by us which were presented in the previous sections.

4.5.1 Basic structure of the model

The basic elements of a model describing the single-tone response patterns of a single auditory-nerve fiber are depicted in Figure 4.21. These elements are a linear filter, a nonlinear

Figure 4.21. Block Diagram of the Structure of a Model of the Responses to a Single Tone

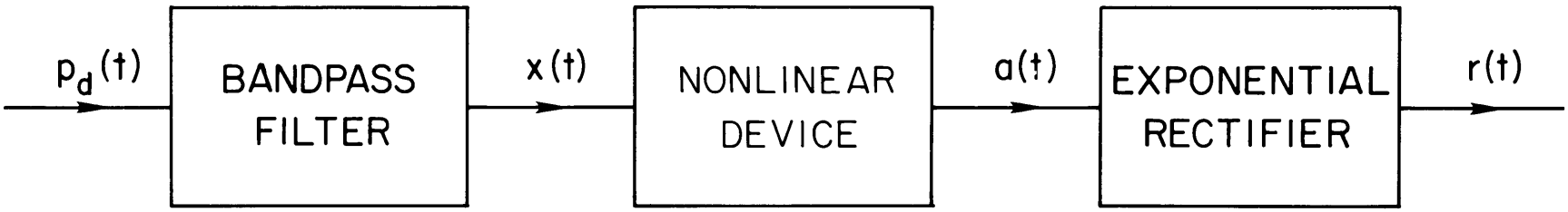
The general structure of a model of single-tone response patterns consists of a cascade of three basic elements. The model is intended to related the rate of discharge r(t) of a single auditory-nerve fiber to the sound pressure $p_d(t)$ at the tympanic membrane. The first element of this model is a linear, time-invariant, bandpass filter representing the tuning properties of single-tone response patterns. The output of this filter, x(t), is then a sinusoid; this signal is passed through a nonlinear device whose output a(t) is written:

$$a(t) = a_o + a_1 \cos 2\pi(ft+\theta_1) + a_2 \cos 2\pi(2ft+\theta_2)$$

In general, the five parameters of the signal a(t) are functions of the level and frequency of x(t). The rate of discharge is related to a(t) by an exponential rectifier:

$$r(t) = \exp\{a(t)\}$$

The resulting expression for the rate of discharge corresponds to the exponential approximation used to describe measured period histograms.


$p_d(t)$
BANDPASS FILTER
$x(t)$
NONLINEAR DEVICE
$a(t)$
EXPONENTIAL RECTIFIER
$r(t)$

device, and an exponential rectifier. The responses of individual auditory-nerve fibers do not interact in this scheme.

The first element of the model is a linear, time-invariant, bandpass filter; it is intended to represent the tuned properties of single-tone responses (i.e., the tuning curve). Conceptually, this filter combines the frequency variations of the middle-ear transfer function (Guinan and Peake, 1967) and of "cochlear tuning." The remaining elements of the model may depend upon frequency, but these elements do not depend upon the CF of the fiber.

If x(t), the output of the bandpass filter, is written

$$x(t) = X \cos 2\pi ft, \tag{4.6}$$

the resulting output of the next stage in the cascade, a(t), is given by

$$a(t) = a_o + a_1 \cos 2\pi(ft+\theta_1) + a_2 \cos 2\pi(2ft+\theta_2) \tag{4.7}$$

where

a_o, a_1, a_2, θ_2 are functions of X and f.

θ_1 is a function of f only.

The output of this element therefore contains a time-independent component and time-varying components at the first and second harmonics of the input sinusoid x(t). This element is simply

denoted as a "nonlinear device." The parameters of a(t) are identical to the parameters of the exponential approximation (equation 4.4). Consequently, we shall concentrate upon a more detailed specification of the "nonlinear device" in the following sections.

The rate of discharge r(t) is determined by passing a(t) through an exponential rectifier

$$r(t) = \exp\{a(t)\} \tag{4.8}$$

This variable serves as the output of the model.

A spike train is not produced by this model. The statistics of measured response patterns do not affect the waveforms of period histograms (Gray, 1966). Consequently, a more detailed model containing a point process description of auditory-nerve fiber discharge patterns is not necessary in order to describe the discharge rates of single-tone response patterns.

4.5.2 Specification of the nonlinear device

The data presented in section 4.4 bear upon the relationship of a(t) to x(t), the output and input respectively of the "nonlinear device" depicted in Figure 4.21. The details of this element of the model are shown in Figure 4.22. The "nonlinear device" is broken down into three branches, each branch corresponding to one of the components in a(t) (equation 4.7). The amplitudes and phases of each component are determined by the elements

Figure 4.22. Block Diagram of the Model Describing Single-Tone Response Patterns

The block diagram of the entire model for the description of single-tone response patterns is depicted. The model applies when the input signal $p_d(t)$ is a sinusoid. The variables and systems comprising the model are described below:

p_d is the input variable representing the sound pressure at the tympanic membrane

T(f) is the transfer function of a linear, time-invariant bandpass filter. This stage represents the tuning properties of single-tone response patterns.

x(t) is the output of the bandpass filter T(f). For purposes of specifying the behavior of the components of the model, x(t) is written

$$x(t) = X \cos 2\pi ft$$

$AGC_o[x(t)]$ is an automatic gain control producing an output b_o which does not vary with time.

The characteristics of this system are given by:

$$b_o = \ln r_{sp} + \frac{\sqrt{2}\,\mathrm{RMS}\,[x(t)]}{2 + \sqrt{2}\,\mathrm{RMS}[x(t)]} \cdot [b_o(\infty) - \ln r_{sp}]$$

where:

RMS [x(t)] denotes the rms value of the waveform x(t). Given the specification of x(t) as a sinusoid:

$$\mathrm{RMS}\,[x(t)] = \frac{X}{\sqrt{2}}$$

(4.22) r_{sp} denotes the spontaneous discharge rate of the output of the entire model.

$b_o(\infty)$ denotes the value approached by b_o as X becomes large. This term depends upon the frequency f according to the empirically-determined formula:

$$b_o(\infty) = \frac{5 \cdot [0.8\,f + 0.8\,\exp(f)]}{[2.37 + 0.8\,\exp(f)]}, \qquad f \text{ in kHz}$$

This description of $b_o(\infty)$ is intended to approximate the main trend of the data given in Figure 4.18.

$AGC_1[x(t)]$ represents an automatic gain control having output $b_1(t)$. Its characteristics are given by:

$$b_1(t) = \frac{6.5 \cdot x(t)}{1 + \sqrt{2}\,\mathrm{RMS}\,[x(t)]}$$

$x^2(t)$ represents a device exhibiting a square-law behavior for time-varying x(t). The output $y_2(t)$ does not contain zero-frequency components. With the specification of x(t) given above:

$$y_2(t) = X^2 \cos 2\pi(2ft)$$

$H_A(f)$ represents the transfer function of a first-order, linear, time-invariant, allpass filter. $H_A(f)$ is expressed by:

$$H_A(2f) = \frac{j\left(\frac{2f}{f_o}\right) - 1}{j\left(\frac{2f}{f_o}\right) + 1}, \qquad f \text{ in kHz}$$

where:

(4.22) f_o is a parameter which may depend upon the level of $y_2(t)$.

$$f_o = \begin{cases} 1.4 \text{ kHz} & f > 0.5 \text{ kHz} \\ 1.4 + \dfrac{0.8 \cdot \sqrt{2}\, \text{RMS}\,[y_2(t)]}{(4)^2 + \sqrt{2}\,\text{RMS}\,[y_2(t)]} & f \leq 0.5 \text{ kHz} \end{cases}$$

The output of this allpass filter is denoted by $z_2(t)$. Consequently:

$$z_2(t) = X^2 \cos 2\pi(2ft + \angle H_A(2f))$$

$AGC_2[z_2(t)]$ represents an automatic gain control with output $b_2(t)$. This AGC is specified by:

$$b_2(t) = \frac{2.8 \cdot z_2(t)}{(4)^2 + (\sqrt{2} \cdot \text{RMS}[z_2(t)])^{1/2} + \sqrt{2}\,\text{RMS}[z_2(t)]}$$

$H_L(f)$ is the transfer function of a linear, time-invariant, lowpass filter. Its input consists of the sum of b_o, $b_1(t)$, and $b_2(t)$. Its output is labeled $a(t)$. $H_L(f)$ is given by:

$$H_L(f) = \frac{1}{\left(1 + j\,\frac{f}{0.63}\right) \cdot \left(1 + j\,\frac{f}{3}\right)}, \qquad f \text{ in kHz}$$

$a(t)$ is the output of the lowpass filter and is written:

$$a(t) = a_o + a_1 \cos 2\pi(ft + \theta_1) + a_2 \cos 2\pi(2ft + \theta_2)$$

$\exp[a(t)]$ denotes an exponential rectifier with input signal $a(t)$. The output $r(t)$ represents the rate of discharge.

$r(t) = \exp[a(t)]$.

(4.22)

In summary, the parameters of a(t) are given by:

$$a_o = \ln r_{sp} + \left(\frac{X}{2+X}\right) \cdot \left\{ \frac{5 \cdot [0.8f + 0.8\exp(f)]}{[2.37 + 0.8\exp(f)]} - \ln r_{sp} \right\}$$

$$a_1 = \frac{6.5X}{1+X} \cdot \frac{1}{\left[1 + \left(\frac{f}{0.63}\right)^2\right]^{1/2} \left[1 + \left(\frac{f}{3}\right)^2\right]^{1/2}}$$

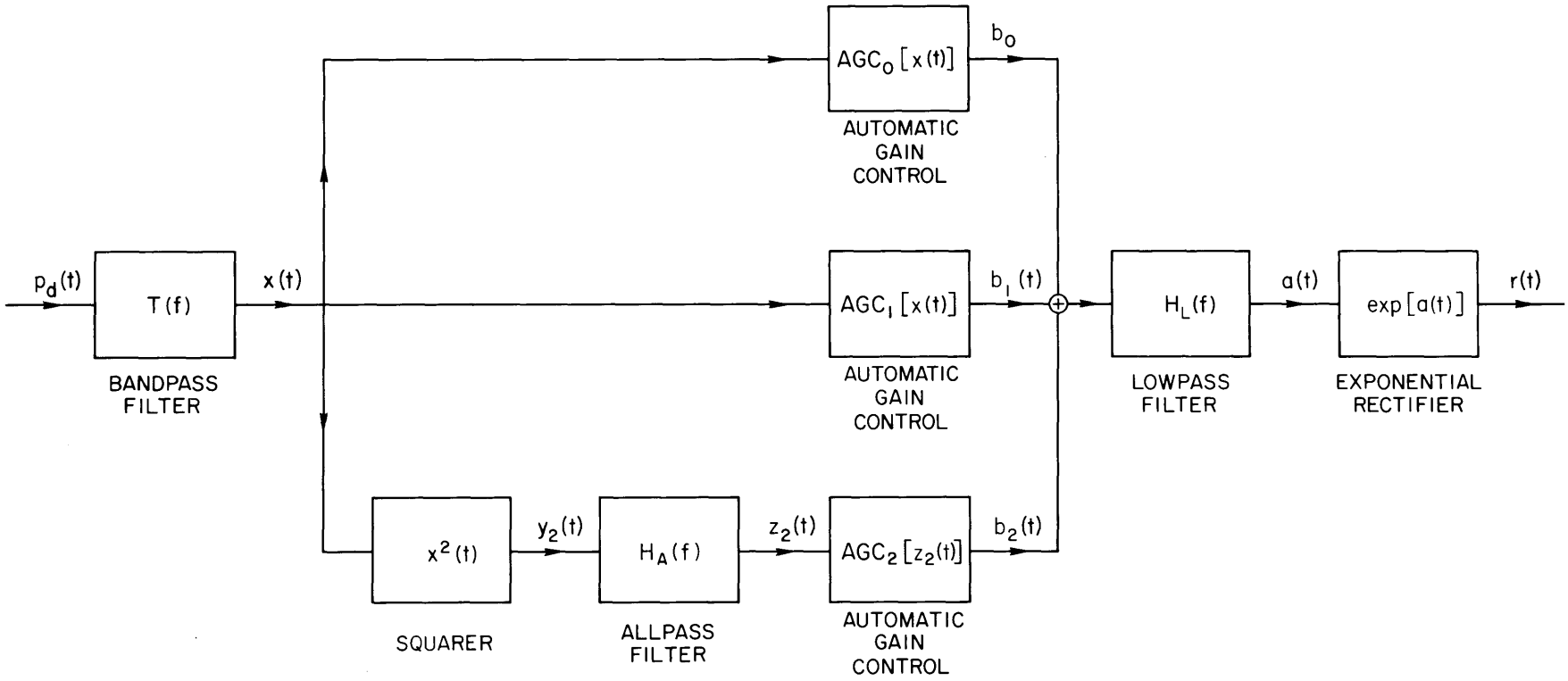
$$\theta_1 = -\frac{1}{2\pi}\left[\tan^{-1}\left(\frac{f}{0.63}\right) + \tan^{-1}\left(\frac{f}{3}\right)\right]$$

$$\theta_2 = .5 - \frac{1}{\pi}\tan^{-1}\left(\frac{f}{f_o}\right) - \frac{1}{2\pi}\left[\tan^{-1}\left(\frac{2f}{0.63}\right) + \tan^{-1}\left(\frac{2f}{3}\right)\right]$$

This specification of a(t) presumes that the phase of x(t) is zero. If x(t) has phase β:

$$x(t) = X \cos 2\pi(ft+\beta)$$

then a_o, a_1, and a_2 are unchanged while θ_1 and θ_2 are changed by the addition of β and 2β respectively.


$p_d(t)$
T(f)
BANDPASS FILTER
x(t)
$AGC_0[x(t)]$
AUTOMATIC GAIN CONTROL
b_0
$AGC_1[x(t)]$
AUTOMATIC GAIN CONTROL
$b_1(t)$
$x^2(t)$
SQUARER
$y_2(t)$
$H_A(f)$
ALLPASS FILTER
$z_2(t)$
$AGC_2[z_2(t)]$
AUTOMATIC GAIN CONTROL
$b_2(t)$
$H_L(f)$
LOWPASS FILTER
a(t)
exp[a(t)]
EXPONENTIAL RECTIFIER
r(t)

in each branch.

The outputs of these branches are added together and passed through a linear, time-invariant lowpass filter having a transfer function denoted by $H_L(f)$. This filter is intended to describe the decrease of the a_1 and a_2 level-maxima as the stimulus frequency is increased (Figure 4.14a, b). As shown in Figure 4.14c, the dependence of these level-maxima upon frequency is nearly indentical if they are plotted versus the frequency of their respective components. This result implies that their dependence upon frequency may be described by the passage of first and second harmonic components through the same lowpass filter. The transfer function $H_L(f)$ is given by

$$H_L(f) = \frac{1}{1 + j\left(\frac{f}{0.63}\right)} \cdot \frac{1}{1 + j\left(\frac{f}{3}\right)} \qquad (4.9)$$

where

f is frequency expressed in kHz.

The dependence of the magnitude of $H_L(f)$ on frequency is compared with the data in Figure 4.14c. Note that time-independent signals (f = 0) are not affected by this lowpass filter.

The amplitudes a_o, a_1, and a_2 are not proportional to the level of the stimulus. Rather, as the stimulus level is increased, they tend to saturate at some value, maintaining this value as level

is increased further. In terms of the model, these amplitudes are saturating functions of the amplitude of x(t), X. A device which modifies the amplitude of its input without affecting the waveform of its input is termed an automatic gain control or AGC. The data illustrated in Figure 4.12 imply that each amplitude parameter is related to X by an AGC independent of the other parameters; for example, a_2 increases in the range of levels where a_1 equals its level-maximum. Consequently, each branch in Figure 4.22 contains an AGC. The general form of the AGCs used in this model is given by

$$\mathrm{AGC}[g(t)] = \frac{K_1 g(t)}{K_2 + F[\mathrm{RMS}\{g(t)\}]} \tag{4.10}$$

where

$g(t)$	is the input signal to the AGC.
K_1, K_2	are constants.
$\mathrm{RMS}\{g(t)\}$	denotes the root-mean-square value of g(t). When $g(t) = G \cos 2\pi ft$, then $\mathrm{RMS}\{g(t)\} = G/\sqrt{2}$.
$F[G/\sqrt{2}]$	is a function which is zero for G = 0 and approaches a linear behavior as G becomes large (i.e., $F[G/\sqrt{2}] \propto G$ for large G).

Suppose $g(t) = G \cos 2\pi ft$. If G is small so that $F[G/\sqrt{2}] < K_2$, then $\mathrm{AGC}[g(t)] = \frac{K_1}{K_2} \cdot g(t)$. As the amplitude of g(t) becomes larger so that $F[G/\sqrt{2}] > K_2$, then $\mathrm{AGC}[g(t)] \propto \frac{K_1 g(t)}{G} \propto K_1 \cos 2\pi ft$; in

other words, the amplitude of the output of the AGC becomes independent of the input amplitude. The constant K_1 denotes the amplitude of the output and K_2 denotes the level at which the output changes from a linear amplitude growth to a saturation of the amplitude. The detailed specification of the AGCs in each branch is given in the caption of Figure 4.22.

The signal a(t) contains a component at twice the frequency of the input. This component is assumed to be generated by a square-law device, depicted in the lower branch of Figure 4.22. The amplitude of the output of this square-law device is proportional to the square of its input amplitude. This behavior is reflected in a comparison of the amplitudes a_1 and a_2 lying below their level-maxima (Figure 4.13). As shown previously, when an AGC is operating in its linear region, its output amplitude is proportional to its input amplitude. Therefore, in this region, a_1 is proportional to X, the amplitude of x(t), and a_2 is proportional to the amplitude of the output of the square-law device. A comparison of a_1 and a_2 in this operating region of an AGC is equivalent to a comparison of X and the output of the squarer. Therefore, the square-law relationship of a_1 and a_2 is consistent with deriving the second-harmonic component from a square-law device. Note that this device is not a true squarer; any time-dependent components in the result of squaring its input are not present in its output.

Figure 4.23. Relative Phase Characteristics of the First and Second Harmonics Given a Lowpass-Filter Description of Their Amplitude Characteristics

In describing the frequency-dependent characteristics of the level-maxima of a_1 and a_2 by a single realizable, linear, time-invariant, lowpass filter, the relative phase characteristics of the first and second harmonics are changed. The phase estimates θ_1 and θ_2 are the measured phase angles of the first and second harmonics respectively at the output of the lowpass filter. The average behavior of the relative phase $\theta_2 - 2\theta_1$ with frequency is shown in the lower panel of Figure 4.16. Define θ_1' and θ_2' to be the phase angles of the first and second harmonics respectively at the input of the lowpass filter. Allowing $H_L(f)$ to denote the transfer function of this filter, the input and output phases are related by:

$$\theta_1 = \theta_1' + \angle H_L(f)$$

$$\theta_2 = \theta_2' + \angle H_L(2f)$$

Consequently, the relative phase angle of the first and second harmonics at the input of the lowpass filter $\theta_2' - 2\theta_1'$ is given by:

(4.23)

$$\theta_2' - 2\theta_1' = \theta_2 - 2\theta_1 + \frac{2}{2\pi}\left[-\tan^{-1}\left(\frac{f}{0.63}\right) - \tan^{-1}\left(\frac{f}{3}\right)\right] - \frac{1}{2\pi}\left[-\tan^{-1}\left(\frac{2\cdot f}{0.63}\right) - \tan^{-1}\left(\frac{2\cdot f}{3}\right)\right]$$

where

f is the tone frequency in kHz.

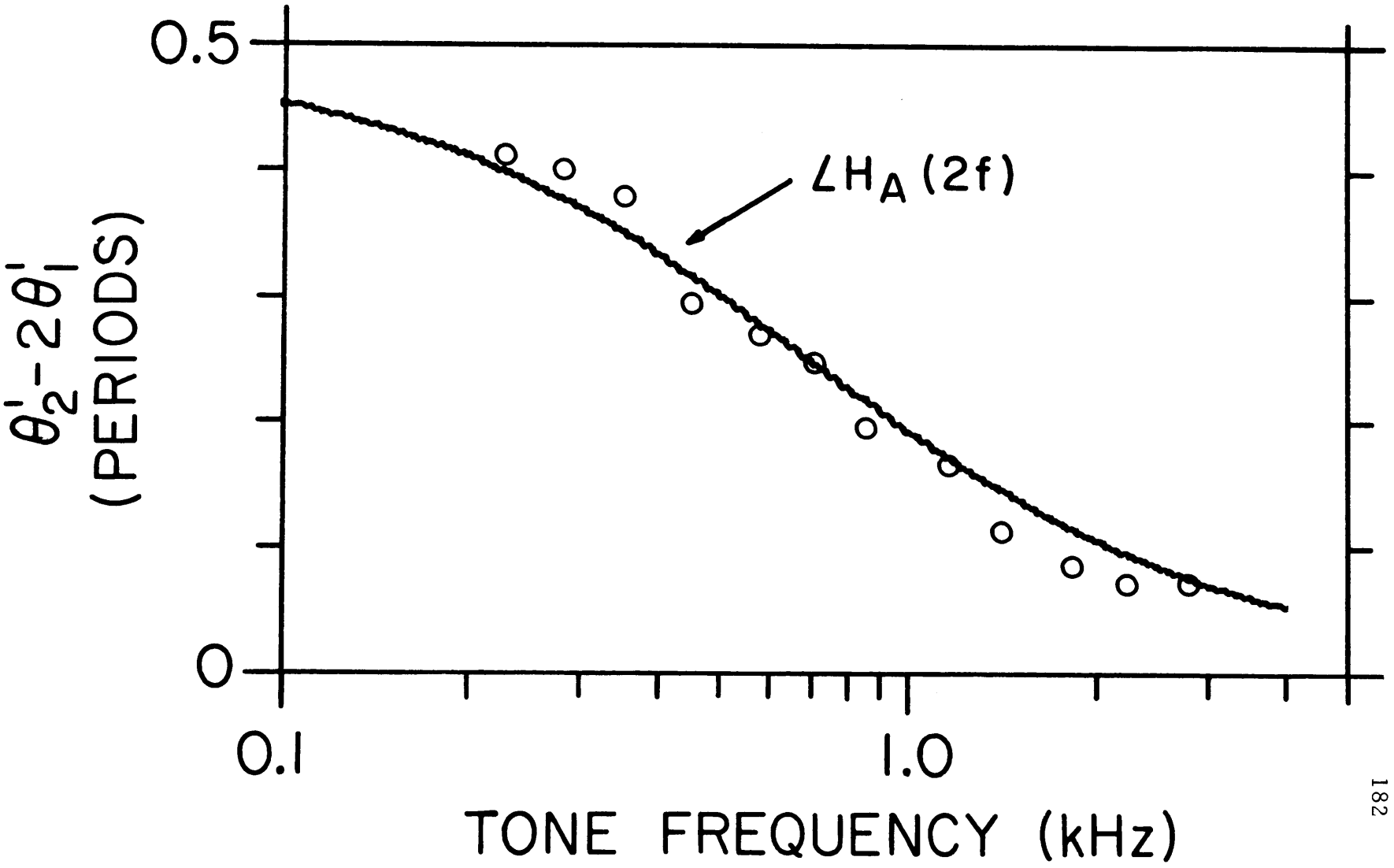
The data points displayed in this figure were derived from the third-octave averages of $\theta_2 - 2\theta_1$ depicted in Figure 4.17 using the above equation. The solid line is an approximation to these derived data points.

This approximation is given by the phase of the first-order allpass filter $H_A(2f)$.

$$H_A(2f) = \frac{j\frac{2f}{1.4} - 1}{j\frac{2f}{1.4} + 1}$$

$$\angle H_A(2f) = 0.5 - \frac{1}{\pi}\tan^{-1}\left(\frac{2f}{1.4}\right)$$

Positive phase indicates lead of the second harmonic relative to the first harmonic.


0.5
0
0.1
1.0
$\theta_2'-2\theta_1'$ (PERIODS)
$\angle H_A(2f)$
TONE FREQUENCY (kHz)

The manner in which the amplitude parameters a_1 and a_2 are determined by the model has been presented; we turn now to an analysis of the phase effects present in the model. The measurements of the relative phase $\theta_2-2\theta_1$ are shown in Figure 4.16. The lowpass filter $H_L(f)$ applies a phase shift to its input. Consequently, if θ_1' is defined to be the phase of the fundamental at the input to the lowpass filter and if θ_2' is the similarly defined phase of the second harmonic, then we have

$$\theta_1' = \theta_1 - \angle H_L(f) \tag{4.11a}$$

$$\theta_2' = \theta_2 - \angle H_L(2f) \tag{4.11b}$$

where

$\angle H_L(f)$ is the phase angle of the transfer function $H_L(f)$ given by equation 4.9. Therefore

$$\angle H_L(f) = \frac{-1}{2\pi} \tan^{-1}\left(\frac{f}{0.63}\right) - \frac{1}{2\pi} \tan^{-1}\left(\frac{f}{3}\right) \tag{4.12}$$

Given the relative phase $\theta_2-2\theta_1$ depicted in Figure 4.16b, the result of applying the phase shift of equations 4.11a and 4.11b is shown in Figure 4.23. The small phase shift of $\theta_2-2\theta_1$ with frequency (Figure 4.16b) is now reflected in a large phase shift in $\theta_2'-2\theta_1'$. The major portion of this phase shift is due to the lowpass filter. The phase angle $\theta_2'-2\theta_1'$ is interpreted as the phase shift

(as a function of tone frequency rather than component frequency) that is to be applied to the second harmonic before the lowpass filter. This phase shift is well described by the phase angle of a first-order allpass filter* having its frequency parameter f_o equal to 1.4 kHz. This allpass filter, having system function $H_A(f)$, is placed in the branch of the model carrying the second-harmonic component (Figure 4.22).

$$H_A(f) = \frac{j\left(\frac{f}{1.4}\right) - 1}{j\left(\frac{f}{1.4}\right) + 1} \tag{4.13a}$$

$$\left|H_A(f)\right| = 1 \tag{4.13b}$$

$$\angle H_A(f) = .5 - \frac{1}{\pi} \tan^{-1}\left(\frac{f}{1.4}\right) \tag{4.13c}$$

where

f is frequency in kHz.

As noted in Figure 4.15, there can be phase shifts in the relative phase $\theta_2 - 2\theta_1$ as a function of level for stimulus

*An allpass filter does not influence the amplitude of a sinusoidal input but it does introduce a phase shift.

frequencies below 0.5 kHz. This level-dependent phase shift is important enough to warrant inclusion in the model. The allpass filter accounts for the average behavior of $\theta_2 - 2\theta_1$ as a function of frequency. The level-dependent phase shifts of $\theta_2 - 2\theta_1$ are described by allowing the frequency parameter of $H_A(f)$ to vary with level so that $\theta_2 - 2\theta_1$ approaches 0.5 as level is increased. The exact description of the resulting allpass filter is given in the caption of Figure 4.22.

The last aspect of the model to be specified the relationship of a_o to x(t). For the sake of clarity, let $a_o(f,X)$ represent the value of a_o obtained as a function of the frequency f and the amplitude of x(t), X. If X = 0, a_o is given by equation 4.5

$$a_o(f, 0) = \ln r_{sp} \tag{4.14}$$

As indicated in Figure 4.18, the values of a_o obtained at large values of X are a function of frequency only. These relationships can be described by

$$a_o(f,X) = \ln r_{sp} + AGC_o(X) \cdot [a_o(f, \infty) - \ln r_{sp}] \tag{4.15}$$

where

$AGC_o(X)$ represents an automatic gain control.

r_{sp} is the spontaneous discharge rate.

a_o is the final value of a_o obtained as a function of frequency.

The automatic gain control $AGC_o(X)$ is assumed to be a monotonic function of X. Non-montonic relationships of a_o with levels are obtained at low frequencies (Figure 4.12). Consequently, equation 4.15 will not describe these non-monotonicities. These relationships were not included in the model for the sake of simplicity.

4.5.3 Characteristics of the Model

A model of the response of an auditory-nerve fiber to a single tone has been specified. This model incorporates the exponential approximation of the rate of discharge. The lowpass filter accounts for the decline of the synchronization index with frequency. The automatic gain controls reflect the saturation of the average discharge rate and the synchronization index at high stimulus levels. The square-law device and the allpass filter produce a second-harmonic component with the proper amplitude and phase so as to describe the waveform of measured period histograms.

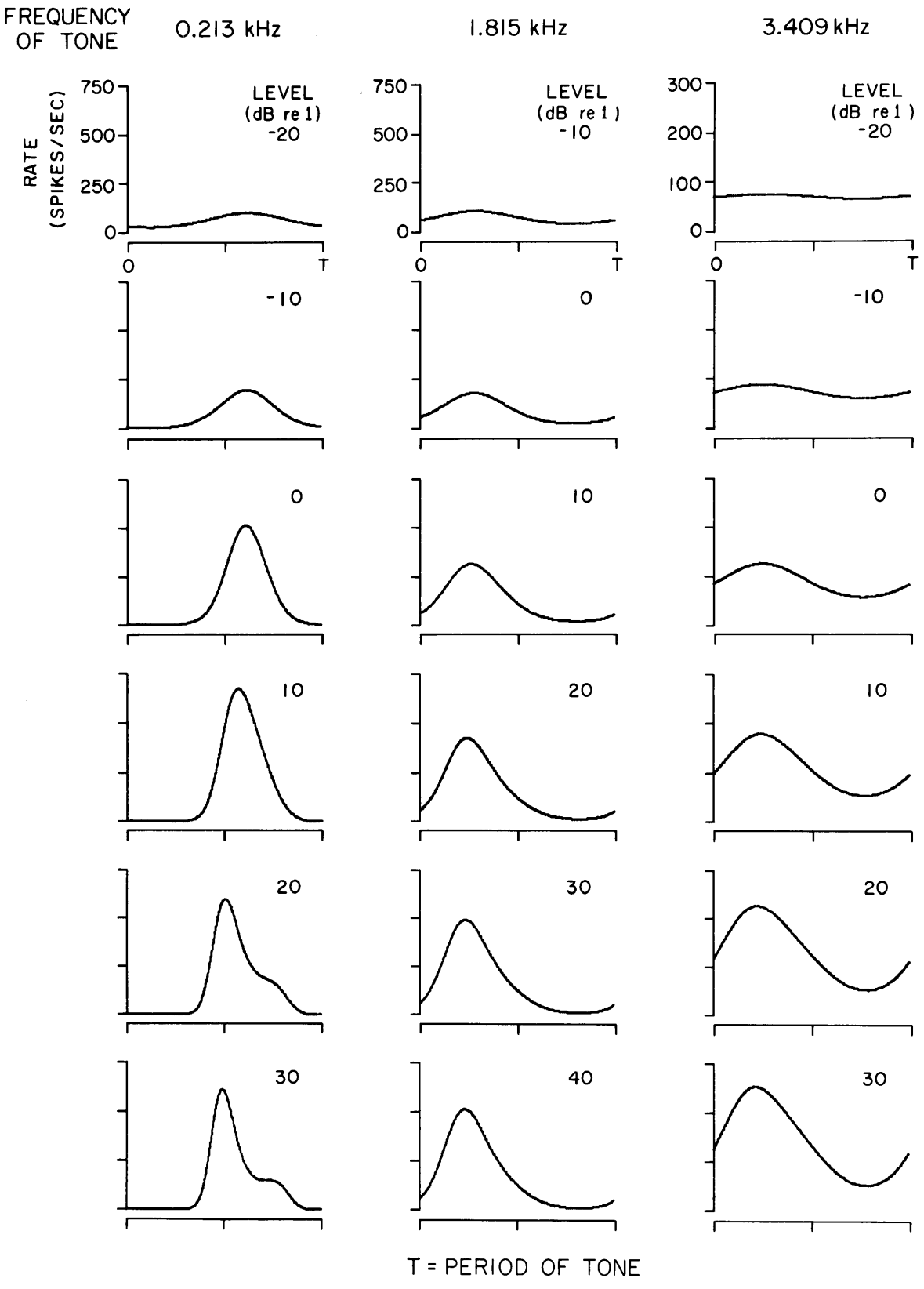
Figure 4.24 displays rates of discharge given by the model for three stimulus frequencies. These frequencies were chosen to match the stimulus frequencies in Figure 4.1. The model therefore can describe the general characteristics of measured period histograms. The average rates and synchronization indices measured from the waveforms computed by the model are displayed

Figure 4.24. Output Waveforms of the Model

The waveforms of r(t) derived from the model are depicted for three different frequencies. These frequencies were chosen to be the same as specified in Figure 4.1. Each column represents the waveforms derived from a level series, the increment being 10 dB. The level is specified as the amplitude X of x(t) expressed in dB re 1. The phase of x(t) was shifted so that the waveforms shown here would have approximately the same phase as the period histograms of Figure 4.1. The spontaneous discharge rate r_{sp} of the model's output was fixed at 70 spikes/sec. The final value of the amplitude parameter a_o was changed for each frequency so that the maximum average rate of discharge would be near the maximum average discharge rates given in Figure 4.3. These values and the values given by the model as specified in Figure 4.22 are:

Frequency (kHz)	$a_o(\infty)$ (Model)	$a_o(\infty)$ (Figure)
0.213	1.73	1.33
1.815	4.37	4.48
3.409	5.07	4.86

The vertical scale of each waveform is expressed in spikes/sec. and the horizontal scale represents one period of the sinusoidal stimulus (T).

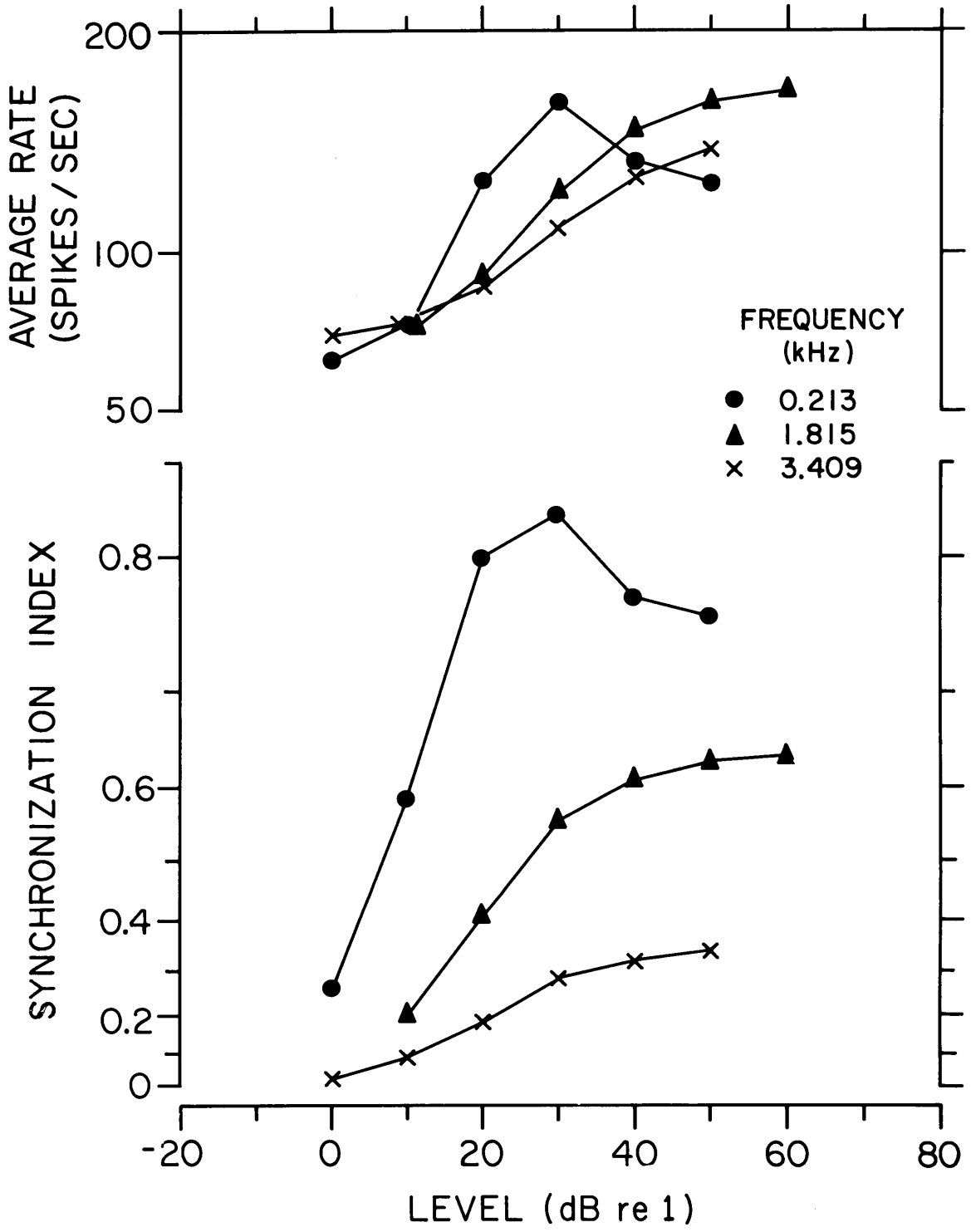

FREQUENCY OF TONE
0.213 kHz
1.815 kHz
3.409 kHz
RATE (SPIKES/SEC)
750
500
250
0
300
200
100
LEVEL (dB re 1)
-20
-10
0
10
20
30
40
T
T = PERIOD OF TONE

in Figure 4.25. The synchronization indices computed from the model compare well with those displayed in Figure 4.3. The synchronization indices of the model tend to not increase with level as fast as the measured values. The average rates obtained from the model also tend to increase at a lower rate than the data of Figure 4.3. Note especially that the average rate decreases at high levels in the 0.213 kHz level series computed from the model. This result is a consequence of assuming a monotonic AGC in the specification of a_o (equation 4.15). The average rates continue to increase in regions where the level-maxima of the synchronization indices are attained. Generally speaking, the synchronization indices are more accurately described by the model than are the average discharge rates.

"Peak-splitting" response patterns can be incorporated into the model by allowing the amplitude of the second harmonic to attain larger values. This can be accomplished by increasing the constant in the numerator of the AGC in the branch carrying the second-harmonic component.

Figure 4.25. Average Discharge Rate and Synchronization Index of the Output of the Model: Variation with Level

The average discharge rates and synchronization indices were measured from the waveforms depicted in Figure 4.24. The level is specified as the amplitude X of x(t) and is expressed in dB re 1.


200
100
50
AVERAGE RATE
(SPIKES / SEC)
FREQUENCY
(kHz)
0.213
1.815
3.409
0.8
0.6
0.4
0.2
0
SYNCHRONIZATION INDEX
-20
0
20
40
60
80
LEVEL (dB re 1)

4.6 Summary

The average rate and the synchronization index were measured from the responses of single auditory-nerve fibers to single tones. The relation of these measures to stimulus parameters was described. These response measures are not simple functions of each other; either measure was shown to capable of changing while the other remained constant. This result implied that these responses must, in part, depend upon different variables in a model of single-tone response patterns.

It was found that the period histograms computed from the response pattern to a single tone could be fit by an expression of the form

$$r(t) = \exp\{a_o + a_1 \cos 2\pi(ft+\theta_1) + a_2 \cos 2\pi(2ft+\theta_2)\} \qquad (4.4)$$

This expression consists of the sum of a fundamental term at the frequency of the tone, a second-harmonic term, and a time-independent term which all form the argument of an exponential. The aspect of the response synchronized to the stimulus was described by the time-varying components of this expression. The time-independent term (a_o) varied only the average rate of discharge. The dependence of the amplitudes and phases of these components on stimulus parameters was described. In terms of the exponential approximation, the amplitude a_2 became

the dominant term as the frequency of the stimulus was increased. This result corresponded to the lack of a synchronized response at high frequencies while an average rate response remained. The presence of a second-harmonic component in the exponential approximation should contain a square-law device.

Based upon the exponential approximation, a mathematical model was formulated to describe the dependence of the rate of discharge upon stimulus parameters. This model characterized the dependence of the parameters of the exponential approximation on the single-tone stimulus.

CHAPTER V

SYNCHRONIZED RESPONSES TO TWO-TONE STIMULI

A model has been derived which can predict the response pattern of an auditory-nerve fiber to a single tone. To explore the generality of this model, we shall compare the responses of the model with the responses of auditory-nerve fibers for two-tone stimuli. In this chapter, we shall describe some aspects of the synchronized response of single fibers to two tone and compare them with the output of the model.

5.1 Introduction

5.1.1 Two-Tone Responses

In response to a two-tone stimulus, an auditory-nerve fiber can exhibit a variety of response patterns. Generally speaking, two categories of two-tone response patterns can be defined: 1) an inhibition (suppression) or an increase of the response to one tone by the presence of the second tone and 2) a generation of frequency components in the rate of discharge not present in the stimulus. This categorization of the data is not meant to imply that these two phenomena are due to different mechanisms.

The average rate of discharge evoked by a tone whose frequency is near the CF of the fiber may be decreased by the addition of a second tone to the stimulus. This phenomenon has been termed "two-tone inhibition" (Sachs and Kiang, 1968); for the

purposes of this thesis, the term two-tone rate suppression is used. The second tone must lie in restricted amplitude and frequency ranges in order for two-tone rate suppression to occur. There are two, distinct rate-suppression regions: one for frequencies less than the CF of the unit, the other lying above CF. The rate-suppression regions are intimately related to the tuning curve of the fiber. Each region overlaps the tuning curve; part of the region lies in the response area, the remainder outside. Consequently, a second tone not by itself evoking an average-rate response may suppress the average-rate response to the CF tone. When two-tone rate suppression occurs, the average rate of discharge cannot be maintained below the spontaneous rate.

The synchronized response to a tone may be suppressed by the additional presentation of a second tone (Hind et al., 1967). This two-tone synchrony suppression does not seem to occur in restricted suppression regions as in two-tone rate suppression. Rather, any tone that can evoke a response in a unit can suppress the synchronized response to another tone (Rose et al., 1974).* Hind (1970) suggested that two-tone rate suppression and two-tone synchrony suppression were due to a common mechanism.

---

*Although the data contained in this reference were obtained from the cochlear nucleus, the authors presumed that their data exhibit characteristics similar to data recorded from auditory-nerve fibers.

A second aspect of the two-tone response pattern is the presence in the response of components having frequencies $pf_1 + qf_2$ where f_1 and f_2 are the stimulus frequencies. The distortion component $2f_1 - f_2$ ($f_1 < f_2 < 2f_1$) has been studied more carefully than have other components. Recordings from single auditory-nerve fibers indicate the presence of a response to the combination frequency $2f_1 - f_2$ (Goldstein and Kiang, 1968; Pfeiffer et al., 1974). Goldstein and Kiang (1968) report a response to $2f_1 - f_2$ which can occur even if neither tone when presented alone can evoke a response. This response can be cancelled by adding to the stimulus a tone having the frequency $2f_1 - f_2$ and the proper amplitude and phase. Furthermore, one amplitude and phase setting can cancel the response to $2f_1 - f_2$ in fibers having different CFs (Goldstein, 1970). On the other hand, Pfeiffer et al. (1974) report measurements of a synchronized response to the frequency $2f_1 - f_2$ which depend upon the amplitudes of the synchronized responses to the stimulus frequencies rather than upon the levels of the tones. We shall comment upon the nature of these differing response patterns later in this study (section 5.5.2).

Responses synchronized to the frequencies $f_1 + f_2$ (the sum frequency) and $f_2 - f_1$ (the difference frequency) have also been measured. If the frequencies of two tones are nearly equal, a response is obtained to the beat frequency $f_2 - f_1$ (Kiang et al., 1965: 111-112). Pfeiffer et al. (1974) report synchronized responses to $f_2 - f_1$ and $f_1 + f_2$. Their data indicate that the synchronized

responses to these frequencies are approximately equal in amplitude and are dependent upon the amplitudes of the synchronized responses to the stimulus frequencies.

5.1.2 Models of two-tone responses

Most models of two-tone response patterns describe only limited aspects of the response. Based upon two-tone rate suppression data, Sachs (1969) derived an analytical specification of the relationship between average discharge rate and the levels and frequencies of the two tones. Goldstein (1972) used Siebert's model (1970) to infer that certain combination-frequency responses (Goldstein and Kiang, 1968) are not due to the exponential transducer.

The model proposed by Pfeiffer and Kim (1973) is designed to describe many features of the two-tone response pattern. The dominant aspect of this model is their characterization of the tuned system (Kim et al., 1973a). The tuned system is represented by a cascade of second-order filters, each having a nonlinear damping term. A rectifier followed by a "model neuron" is placed at the output of the cascade to form the entire model (Kim et al., 1973b). Qualitatively, aspects of two-tone rate suppression and the generation of combination frequencies in the response are found in the output of the model.

5.2 Procedures

Experiments involving a two-tone stimulus were performed on 19 cats resulting in a significant amount of data from 142 units. The basic experimental procedures are described in Chapter III with the following additions. The stimulus frequencies f_1 and f_2 were not harmonically related. At least one of these frequencies was less than or equal to the CF of the fiber. The synchronizations indices to f_1 and f_2 were routinely measured from period histograms synchronized to the stimulus frequencies. Each stimulus presentation lasted 30 seconds with a short interval (~5 seconds) between presentations. The tone frequencies were fixed and the level of one tone held constant while a level series (from low to high levels) of the other tone was performed. The increment between successive levels was usually 10 dB. After a level series was completed, the level of the first tone was increased by 10 dB and another level series was initiated fro the second tone. The level of each tone did not exceed 80 dB SPL.

Because no significant "rest period" was provided between stimulus presentations, the possibility of long-term changes in the synchronized-response characteristics of auditory-nerve fibers must be considered. We examined this question in two cats by presenting a stimulus for an extended period of time and measuring the synchronization index during the presentation interval. The stimulus consisted of a single tone having a level at least 20 dB

greater than the threshold level of the average-rate response (as determined by the tuning-curve measurements). The data used in this study were obtained from 21 units.* All of these data were consistent with the trend of the data to be presented here.

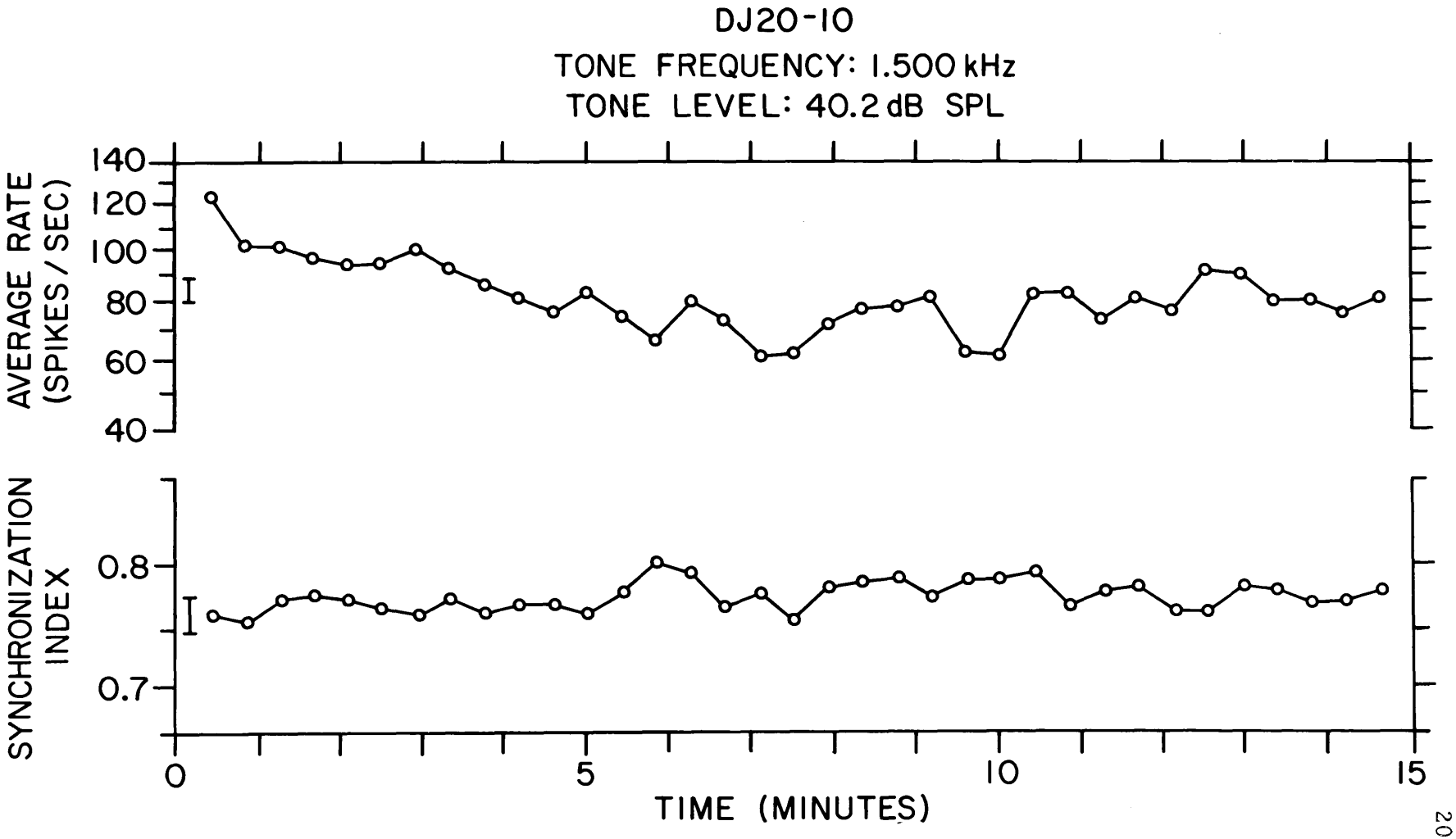
Figure 5.1 presents the results obtained from one auditory-nerve fiber. The average discharge rate decreased steadily during the first five minutes and then became relatively constant. In contrast, the synchronized response, as measured by the synchronization index, remained relatively constant throughout the recording interval. Therefore, the synchronization index does not change with time for stimuli of prolonged duration. Assuming that these results apply for the two-tone stimulus, the stimulus protocol described above allows consistent measurements of the synchronization index but not of the average discharge rate. Consequently, data concerning the average discharge rates measured from two-tone response patterns are not presented.

---

*The responses of these units were recorded for five minutes or longer.

Figure 5.1. Average Discharge Rate and Synchronization Index of Single-Tone Response Patterns: Variation over Extended Periods of Stimulation

At the time origin, a tone was presented as the stimulus. The tone remained on throughout the time interval. Period histograms synchronized to the stimulus were computed from responses over twenty-second intervals every twenty-five seconds. The average rate and the synchronization index were computed from every histogram. The results of these computation are displayed, the data points placed at the end of every twenty-five-second interval. The vertical bars indicate the range of plus and minus two standard deviation of the response measures. The spontaneous discharge rate of the unit was 7.2 spikes/second and the CF of the unit was 1.44 kHz. The threshold level of an average rate response at 1.500 kHz as taken from the measured tuning curve was 3.2 dB SPL.


DJ20-10
TONE FREQUENCY: 1.500 kHz
TONE LEVEL: 40.2 dB SPL
AVERAGE RATE (SPIKES / SEC)
140
120
100
80
60
40
SYNCHRONIZATION INDEX
0.8
0.7
0
5
10
15
TIME (MINUTES)

5.3 Two-Tone Synchrony Suppression

Figure 5.2 illustrates examples of the two-tone response pattern. As shown in section 3.5.3, synchronization of a period histogram to a given frequency removes the synchronized responses to frequencies inharmonic to the synchronizing frequency. Consequently, only the components in the response synchronized to the stimulus frequencies are depicted in the period histograms of Figure 5.2.

In the left panel of this figure, the level of one tone (frequency equalling 0.411 kHz) is held constant. As the level of the second tone (frequency equalling 0.601 kHz) is increased, the synchronized response to the first tone decreases. At the largest second-tone level depicted (23.2 dB), the synchronized response to the first tone is nearly suppressed. The suppression of the synchronized response to one tone by the addition of a second tone to the stimulus is termed two-tone synchrony suppression.

As the second-tone level is increased, the synchronized response to this component increases. Note that a clear synchronized response to the second tone is present at the lowest second-tone level (13.2 dB) for which suppression occurs.

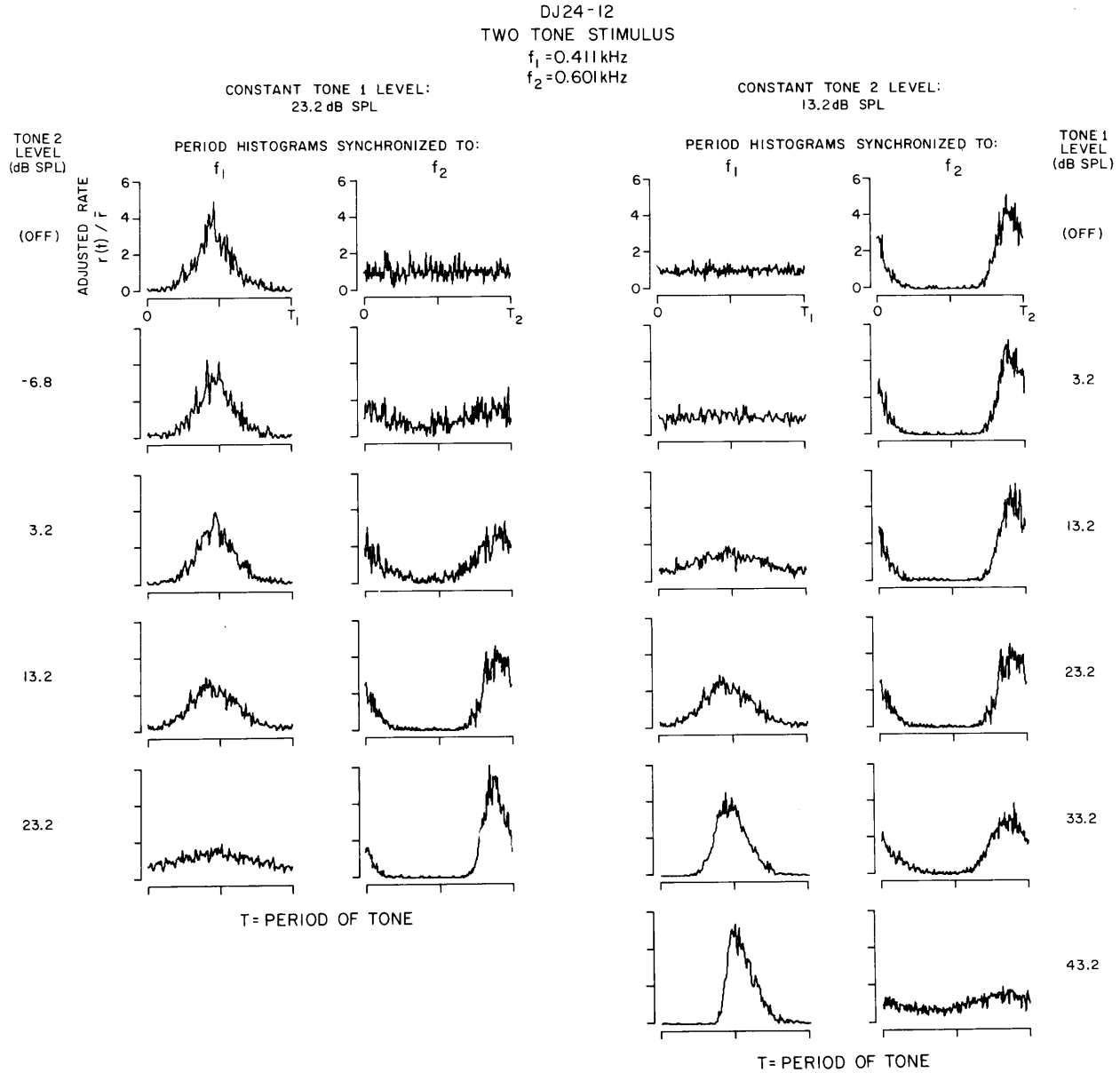
The right panel of Figure 5.2 illustrates the symmetry of two-tone synchrony suppression. Here the level of second tone is held constant while the level of the first tone is increased. Again, the response to the fixed-level tone is suppressed as the level

Figure 5.2. Waveforms of the Rates of Discharge Measured from Two-Tone Response Patterns

In each panel, the stimulus frequencies of both tones and the level of one tone is held constant. The left column of each panel illustrates period histograms synchronized to the frequency f_1 while the right column contains period histograms synchronized to f_2. In the left panel, the level of tone 1 is held constant; the level of tone 2 is increased in 10 dB increments from the top to the bottom of the columns. The top row of histograms were obtained when only tone 1 was presented. Similarly, the level of tone 2 is held constant in the right panel while the level of tone 1 is increased from the top to the bottom of the columns.

The vertical scales are adjusted discharges rates computed by dividing measured period histograms by the average discharge rate. Each horizontal scale represents one period of the synchronizing frequency. The level is specified as the rms sound pressure at the tympanic membrane and is expressed in dB SPL. Each histogram was computed from a thirty second interval.

The spontaneous discharge rate of the unit was 28.0 spikes/second and its CF was 0.58 kHz.


DJ24-12
TWO TONE STIMULUS
f_1 = 0.411 kHz
f_2 = 0.601 kHz
CONSTANT TONE 1 LEVEL:
23.2 dB SPL
CONSTANT TONE 2 LEVEL:
13.2 dB SPL
TONE 2 LEVEL (dB SPL)
PERIOD HISTOGRAMS SYNCHRONIZED TO:
f_1
f_2
ADJUSTED RATE
$r(t) / \bar{r}$
(OFF)
-6.8
3.2
13.2
23.2
33.2
43.2
0
T_1
T_2
T = PERIOD OF TONE
TONE 1 LEVEL (dB SPL)

of the other tone is increased. In summary, the synchronized response to a tone may be suppressed if the level of an additional tone is sufficiently large. Furthermore, the lowest tone level required to suppress the synchronized response to the fixed tone evokes a response to the suppressing tone.

The synchronization indices to the stimulus frequencies are displayed as a function of the level of each tone in Figure 5.3. In the left panel, the synchronization index to the first tone is plotted against the level to the second tone. For example, the synchronization indices depicted for the tone 1 level equalling to 23.2 dB are measured from the period histograms in the left panel of Figure 5.2. As demonstrated in the response waveforms, the synchronization index to tone 1 decreases as the level of the second tone increases. For the tone 1 levels used here, the "knee" of the curves for the synchronization indices to tone 1 occur at approximately the same tone 2 level (≈6-8 dB). This level corresponds to the level of second tone required to just suppress the synchronized response to the first tone. If the second tone is presented alone at this level, the resulting synchronization index to the second tone (the dashed line in the left panel of Figure 5.3) is large (≈0.7). Consequently, the level of second tone required to suppress the first tone evokes a significant response when presented alone. Similar results are obtained if we examine the synchronization index to the second tone, held at constant levels, while the level of the first tone is varied. For low to moderate levels of

Figure 5.3. Synchronization Indices Computed from a Two-Tone Response Pattern

Synchronization indices computed from a two-tone response pattern are displayed as a function of the levels of the two tones. The circles indicate the synchronization indices to f_1 and the triangles indicate the synchronization indices to f_2. Solid lines are used to indicate measurements made when both tones were present in the stimulus while dashed lines indicate the presence of only one tone.

In the left panel, the synchronization indices to the frequency f_1 are displayed as a function of the level of the second tone. Lines connect data points obtained at a constant level of tone 1. Also displayed in the left panel are the synchronization indices to f_2 computed while tone 2 was presented alone. A similar display is found in the right panel; the roles of tone 1 and tone 2 are reversed.

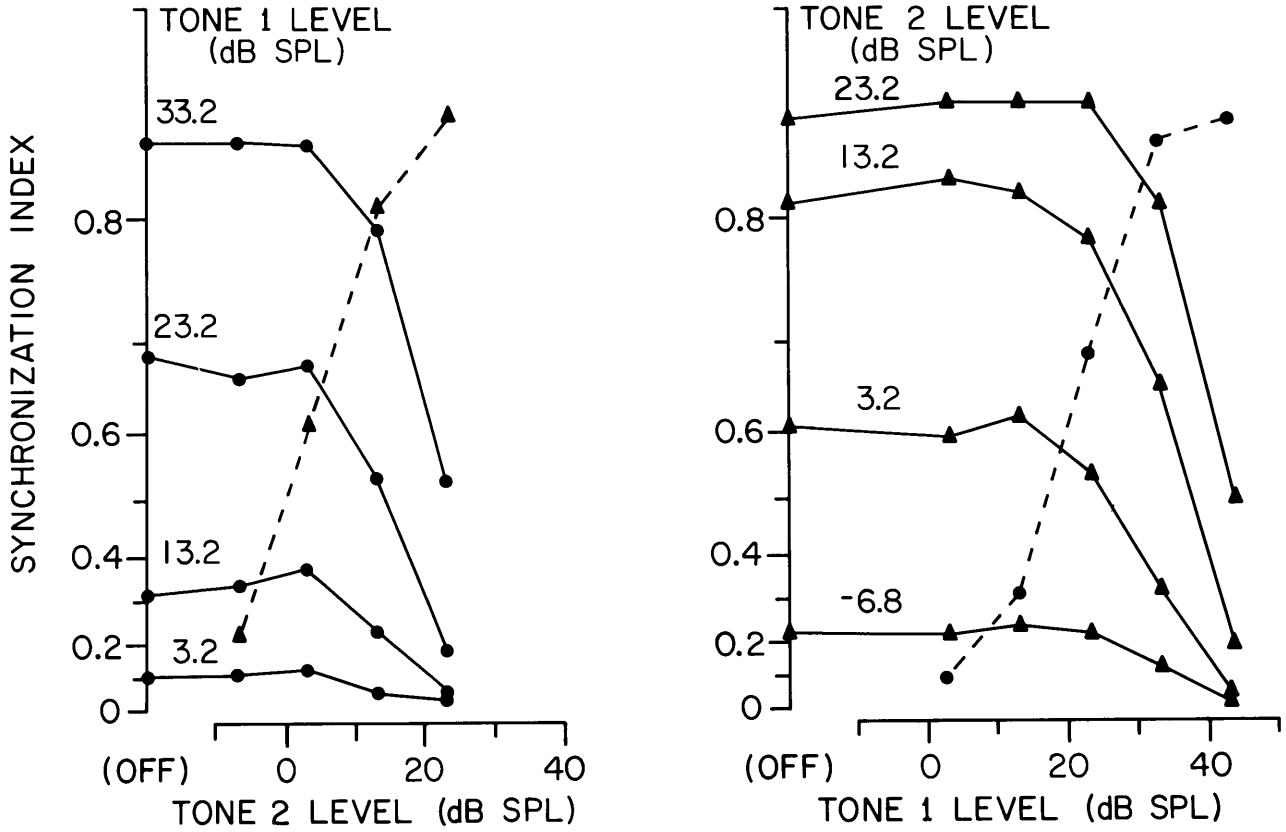
These data were taken from the same unit used in Figure 5.2. Levels are specified as the rms sound pressure at the tympanic membrane.

DJ24-12
TWO-TONE STIMULUS
f_1 = 0.411 kHz
f_2 = 0.601 kHz

SYNCHRONIZATION INDEX TO:

● f_1 BOTH TONES ON ——
▲ f_2 TONE 1 OFF - - - - - -

● f_1 TONE 2 OFF - - - - - -
▲ f_2 BOTH TONES ON ——



the second tone, the "knee" of each curve tends to occur at the same tone 1 level (≈18 dB). However, at larger values of the tone 2 level (e.g., 23.2 dB), this "knee" tends to occur at larger tone 1 levels (≈28 dB).

The results illustrated in Figure 5.3 may be generalized. Two-tone synchrony suppression occurs between two tones as long as each tone, when presented alone, evokes a synchronized response. All of the 142 units studied exhibited two-tone synchrony suppression. These units had CFs which varied over the frequency range 0.19 kHz to 5.44 kHz. Units with larger CFs were not used as the tone levels required to study synchrony suppression would have exceeded 80 dB SPL.

5.4 Frequency Distortion Components

The previous section dealt with the discharges synchronized to the stimulus frequencies in a two-tone response pattern. Figure 5.4 depicts the spectra of the response patterns to a two-tone stimulus. Frequency components other than those harmonic to the stimulus frequencies are readily seen. In particular, the response pattern resulting from the higher stimulus levels (23.2 dB and 13.2 dB for tone 1 and tone 2 respectively) has components of significant amplitudes at the difference frequency $f_2 - f_1$, the sum frequency $f_1 + f_2$, and the combination frequency $2f_1 + f_2$. In this section, we shall focus upon the components at the sum and difference frequencies.

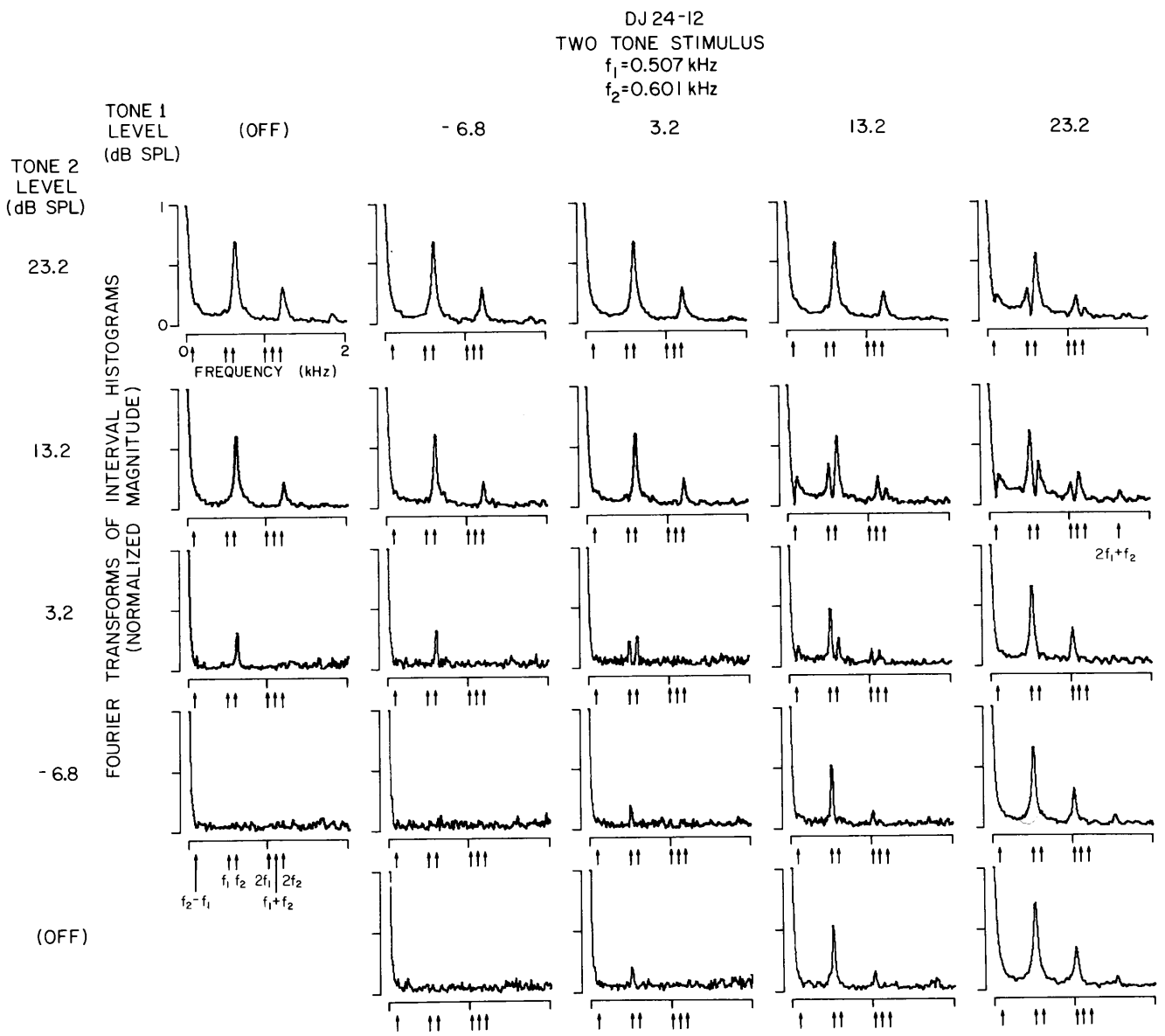
For all of the spectra shown in Figure 5.4, the sum and difference components of the response are smaller in amplitude than either of the components synchronized to the stimulus frequencies. The amplitudes of these distortion components depend upon the amplitudes of the synchronized responses at the rather than the levels of the stimulus tones. Significant amplitudes of the components $f_1 + f_2$ and $f_2 - f_1$ occur only when synchronized responses are present to both of the stimulus frequencies in the two-tone response pattern. For example, consider the column of Figure 5.4 for which the tone 1 level is constant at 13.2 dB. Responses to the sum and difference frequencies can be clearly seen at the 3.2 dB and 13.2 dB levels of tone 2.

Figure 5.4. Normalized Fourier Transforms of Interval Histograms Computed from Two-Tone Response Patterns

Interval histograms were computed from two-tone response patterns over thirty second intervals using a binwidth of 250 μsec. Magnitudes of Fourier transform computed from these histograms were normalized as described in section 3.8.

The level of tone 1 increases along each row from left to right; the level of tone 2 increases from the bottom to the top of each column. The level is specified as the rms sound pressure at the tympanic membrane and is expressed in dB SPL. The frequencies indicated along each horizontal axis are related to the stimulus frequencies as described in the bottom of the first column.

The unit from which these data were recorded is the same unit used in Figures 5.2 and 5.3. However, the frequency of tone 1 is different in this figure.


DJ 24-12
TWO TONE STIMULUS
f_1=0.507 kHz
f_2=0.601 kHz
TONE 1 LEVEL (dB SPL)
(OFF)
- 6.8
3.2
13.2
23.2
TONE 2 LEVEL (dB SPL)
23.2
13.2
3.2
- 6.8
(OFF)
FOURIER TRANSFORMS OF INTERVAL HISTOGRAMS (NORMALIZED MAGNITUDE)
FREQUENCY (kHz)
f_2-f_1
f_1 f_2 $2f_1$ $2f_2$
f_1+f_2
$2f_1+f_2$

However, for the 23.2 dB level of tone 2, the synchronized response at f_1 has been suppressed and no responses appear at $f_1 + f_2$ or $f_2 - f_1$. Furthermore, if the amplitudes of the responses synchronized to f_1 and f_2 are interchanged by suitable changes in the stimulus levels, the amplitudes of the components synchronized to $f_1 + f_2$ and $f_2 - f_1$ remain roughly constant. To be specific, consider the row corresponding to a constant tone 2 level of 13.2 dB. Between the 13.2 dB and 23.2 dB levels of tone 1, the response amplitudes to f_1 and f_2 are approximately interchanged, while the response amplitudes to $f_1 + f_2$ and $f_2 - f_1$ are not changed appreciable. Consequently, the synchronized responses to the sum and difference frequencies as illustrated by these data, depend upon the presence of responses to the stimulus frequencies and furthermore, the amplitudes of the components $f_1 + f_2$ and $f_2 - f_1$ in the response appear to be symmetric functions of the amplitudes of the components f_1 and f_2.

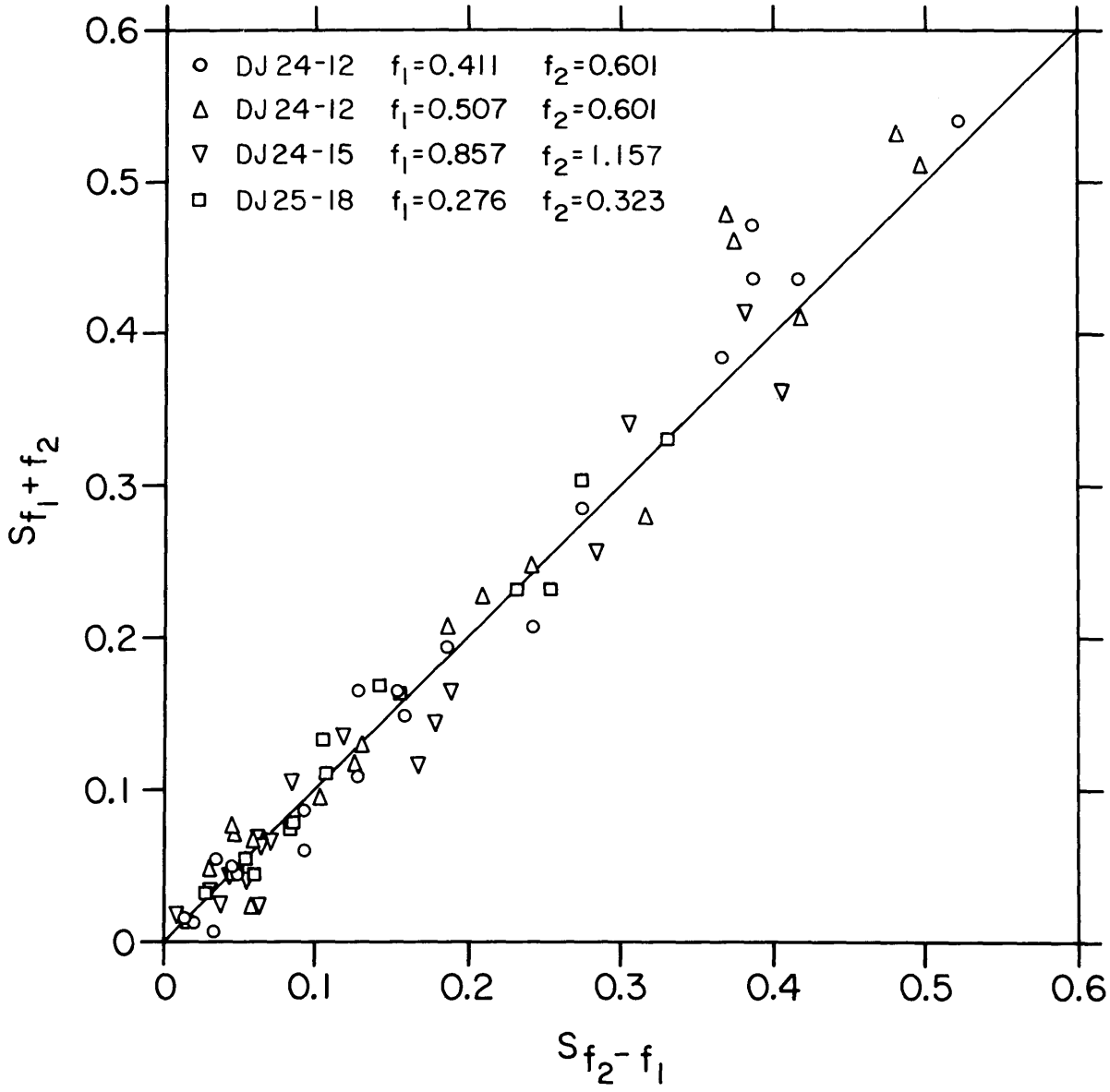
In Figure 5.5, the relative sizes of the synchronization indices measured with respect to $f_1 + f_2$ and $f_2 - f_1$ are compared. The response amplitudes of these components, as measured by the synchronization index, are roughly equal. In one example (unit DJ24-15), the sum frequency is greater than 2 kHz and its amplitude is equal to the amplitude of the difference frequency component, the difference frequency equalling 0.3 kHz. Consequently, responses synchronized to the sum and difference frequencies in a two-tone response pattern are of equal amplitudes for low

Figure 5.5. Relationship of the Amplitudes of Response Components Synchronized to the Sum Frequency and to the Difference Frequencies

Response patterns to a two-tone stimulus having frequencies f_1 and f_2 were recorded from 3 units. Period histograms synchronized to the sum frequency $f_1 + f_2$ and the difference frequency $f_2 - f_1$ were then computed from these response patterns. From these period histograms, synchronization indiceswere computed.

The synchronization indices for the sum- and difference-frequency components contained in the same response pattern are plotted against each other. The line indicates equality between the two measures. The stimulus and unit characteristics were:

Unit No.	r_{sp} (spikes/sec)	CF (kHz)	f_1 (kHz)	f_2 (kHz)	$f_1 + f_2$ (kHz)	$f_2 - f_1$ (kHz)
DJ24-12	28.0	0.58	0.411	0.601	1.012	0.190
DJ24-12	28.0	0.58	0.507	0.601	1.108	0.094
DJ24-15	78.7	1.16	0.857	1.157	2.014	0.300
DJ25-18	74.9	0.47	0.276	0.323	0.599	0.047


DJ 24-12 f_1=0.411 f_2=0.601
DJ 24-12 f_1=0.507 f_2=0.601
DJ24-15 f_1=0.857 f_2=1.157
DJ25-18 f_1=0.276 f_2=0.323
Sf_1+f_2
Sf_2-f_1
0
0.1
0.2
0.3
0.4
0.5
0.6

stimulus level.

The results we have described concerning responses to sum and difference frequencies were obtained in units where large values of the synchronization indices (i. e., greater than about 0. 5) to the stimulus frequencies were obtained at low absolute levels (less than about 40 dB SPL) and, occasionally, at higher levels. Responses were measured which did not conform to the description given above. The data presented in Figure 3. 8 provide an example. Here the difference-frequency component has the largest amplitude in the response and is much larger than the sum-frequency component. These data were measured at larger stimulus levels and the tone frequencies were both greater than the CF of the unit. These responses were not studied in detail.

5.5 Application of the Model to Two-Tone Response Patterns

The model derived in Chapter IV has the capability of describing the response pattern of an auditory-nerve fiber to a single tone having a level less than 80 dB SPL. In particular, waveforms of the rates of discharge and measures of the synchronized responses are well described. In this section, the ability of the model to describe aspects of the two-tone response pattern is assessed.

The model is described in Figure 4.22 for single-tone stimuli. As the model is nonlinear, the model must be defined for two-tone inputs. Assume that x(t) is written

$$x(t) = X_1 \cos 2\pi f_1 t + X_2 \cos 2\pi f_2 t \qquad (5.1)$$

In the upper branch, the relationship between b_o and x(t) for a single sinusoid depends upon the frequency of the sinusoid (equation 4.15). With two-tone stimuli, the relationship between b_o and x(t) is not well defined. As the parameter a_o affects only the average discharge rate and we are concerned with the synchronized response in this chapter, there is no need to define the upper branch of the model for two-tone stimuli. Consequently, the behavior of this branch for two-tone stimuli is left unspecified. The only element in the middle branch of Figure 4.22 is an AGC. The AGCs, as expressed in equation 4.10, are well defined for any signal. Therefore, the operation of the middle branch does

not require any clarification. The first element of the lower branch is the square-law device. Its output for an input consisting of two sinusoids (equation 5.1) is given by

$$y_2(t) = X_1^2 \cos 2\pi(2f_1)t + X_2^2 \cos 2\pi(2f_2t)$$

$$+ 2X_1X_2 \cos 2\pi(f_1+f_2)t + 2X_1X_2 \cos 2\pi(f_1-f_2)t \qquad (5.2)$$

The next element in the lower branch is the allpass filter. In the single-tone model, the parameter f_o of the allpass filter varied with level if the stimulus frequency was less than 0.5 kHz. Consequently, we must define the phase characteristic of the allpass filter when its input, $y_2(t)$, contains four components (equation 5.2). We shall arbitrarily define that the phase shift of each sinusoidal component of the input $y_2(t)$ is applied independently of the frequencies and amplitudes of the other components. Therefore, the level-dependent behavior of the allpass filter is determined independently for each component according to the frequency and level of that component. This specification of the operation of the allpass filter does not affect the results presented here to a great extent. The final element of the bottom branch is an AGC which is well defined for a four-component signal. The remaining elements of the model are well defined.

5.5.1 Two-tone synchrony suppression

The model, as described previously for a two-tone stimulus, demonstrates two-tone synchrony suppression (Figure 5.6). The characteristics of the model, with regard to synchrony suppression, correspond well to those of the data (Figure 5.3). The level of tone necessary to just suppress the synchronized response to a second tone is roughly constant for low to moderate levels of the second tone. As the second tone level becomes large enough to evoke a synchronized response near the level-maximum (e.g., the 20 dB level of tone 2 in the right panel of Figure 5.6), larger levels of the first tone are required to suppress the response synchronized to the second tone. When the first tone is presented alone at the smallest level required for suppression of the response to the first tone (-10 dB in the right panel of Figure 5.6), a significant synchronization index is measured.

In terms of the model, the source of two-tone synchrony suppression is the AGC for the fundamental components. The output of this AGC, $b_1(t)$, is written for an input x(t) given in equation 5.1 as

$$b_1(t) = \frac{X_1}{1 + \left(X_1^2 + X_2^2\right)^{1/2}} \cos 2\pi f_1 t + \frac{X_2}{1 + \left(X_1^2 + X_2^2\right)^{1/2}} \cos 2\pi f_2 t \qquad (5.3)$$

Figure 5.6. Synchronization Indices Measured from the Output of the Model to an Input Consisting of Two Sinusoids

Two sinusoids were used as the input signal x(t) of the model described in Figure 4.22. The circles represent synchronization indices measured with respect to the input frequency f_1 while triangles indicate measurements with respect to the input frequency f_2. Solid lines indicate measurements performed with both tones presented simultaneously and dashed lines indicate measurements performed with a single-tone stimulus. In the left panel, synchronization indices measured with respect to tone 1 are shown as a function of the level of tone 2. The opposite variation is shown in the right panel.

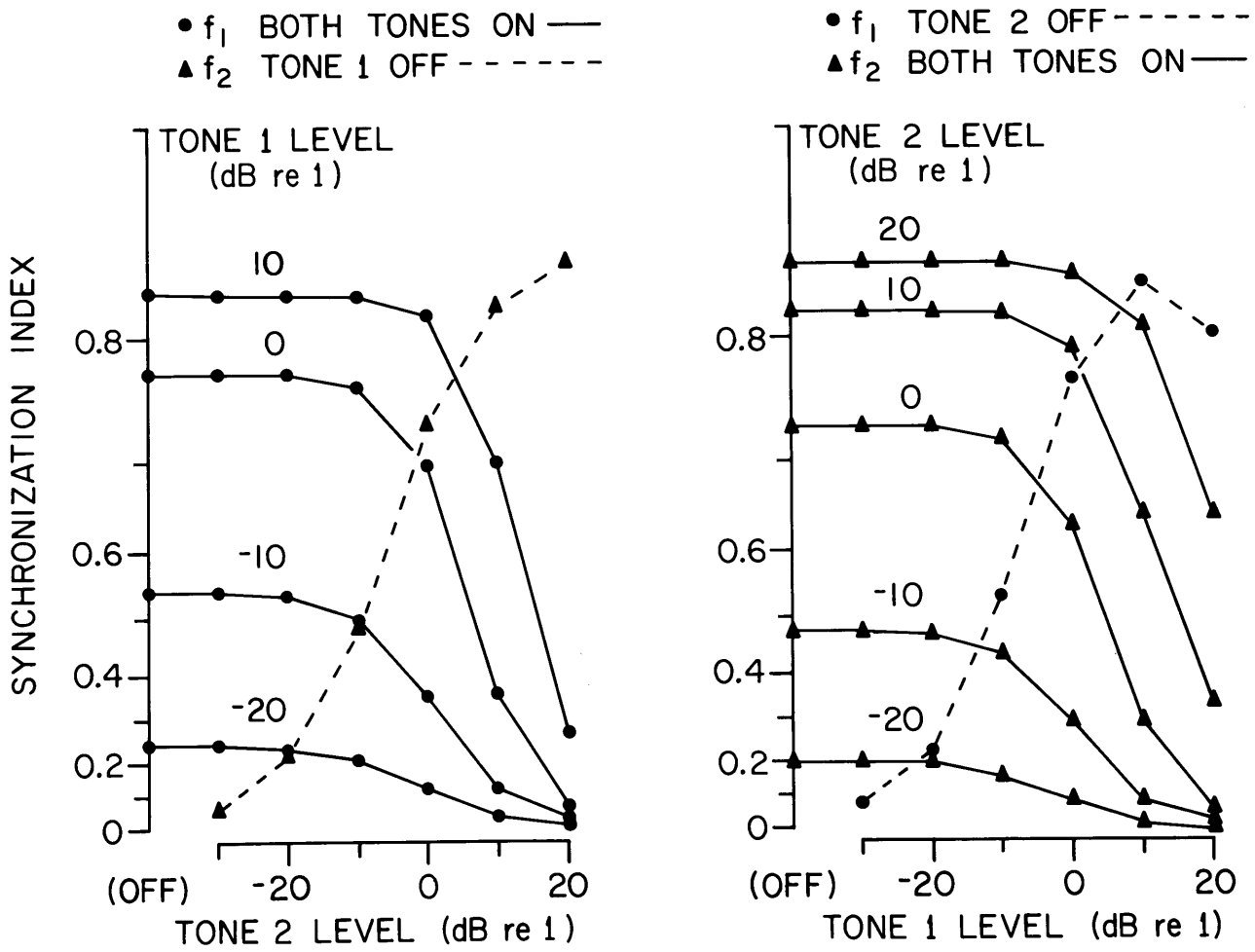
The frequencies chosen for this figure are the same as those used in Figure 5.3. The parameters of the model are as stated in Figure 4.22. Levels are specified as the amplitudes of each component in x(t); these levels are expressed in dB re 1.

MODEL

f_1 = 0.411 kHz

f_2 = 0.601 kHz

SYNCHRONIZATION INDEX TO:



For example, the amplitude of the component at the frequency f_2 is given by

$$\frac{X_2}{1 + \left(X_1^2 + X_2^2\right)^{1/2}} \tag{5.4}$$

If X_2, the level of the second tone, is held constant, then this amplitude will decrease monotonically as X_1 is increased. If X_2 is much less than 1, the level of X_1 which reduces the amplitude of expression 5.4 is constant, approximately equal to 1. This level of X_1 will evoke a significant response when the tone having frequency f_1 is presented alone. As X_2 becomes comparable to 1, X_1 must be greater than X_2 in order to reduce this amplitude (expression 5.4); therefore, larger amplitudes of X_1 are needed for suppression as X_2 becomes large. Suppression of the other component in $b_1(t)$ occurs in a similar fashion; the expression for the amplitudes of each component in equation 5.3 are identical with the roles of X_1 and X_2 interchanged.

Two-tone synchrony suppression is a consequence of a fundamental aspect of the single-tone model: the automatic gain control. In the model, this element results in the saturation of the synchronization index as the stimulus level of a single tone is increased. As seen in Figure 4.3, the saturation of the synchronization index with level is characteristic of the response pattern to a single tone. In terms of the model, this saturation of the response to one tone is

equivalent, when two tones serve as the input, to two-tone synchrony suppression.

5.5.2 Frequency distortion components

As a result of single-tone stimulation, the response pattern of an auditory-nerve fiber contains components not only at the frequency of the stimulus, but also at its harmonics (Figure 4.2). Therefore, significant harmonic distortion is present in the synchronized response pattern to a single tone. This distortion is evident in the shape of the period histogram measured from these responses (Figure 4.1).

In terms of the model, this harmonic distortion is, for the most part, due to the exponential rectifier. The input to the rectifier contains a fundamental component and its second harmonic. In most cases, the fundamental component is much larger in amplitude than is the second harmonic. It is the distortion of the fundamental by the exponential that is the dominant source of the harmonic content in the spectrum of the response.

Many frequency distortion components have been reported (Goldstein and Kiang, 1968; Pfeiffer et al., 1974; section 5.4) in the synchronized response pattern of an auditory-nerve fiber responding to a two-tone stimulus. The purpose of this section is to determine whether any of these measured frequency distortion components can be attributed to the exponential rectification of two sinusoids.

Assume that the time-varying input a(t) of the exponential rectifier can be written

$$a(t) = a_{11} \cos 2\pi(f_1 t+\theta_{11}) + a_{21} \cos 2\pi(f_2 t+\theta_{21}) \qquad (5.5)$$

where

a_{11}, θ_{11} denote the amplitude and phase respectively of the component having frequency f_1

a_{21}, θ_{21} denote the amplitude and phase respectively of the component having frequency f_2

f_1, f_2 are not harmonically-related frequencies.

With respect to the model, this assumption implies that the output of the square-law device does not contribute substantially to the input of the rectifier. The inclusion of the square-law distortion components complicates the analysis with affecting the basic results. The a_o term has been neglected as it does not contribute to the time-varying output of the rectifier.

The output of the rectifier, having an input given by equation 5.5, contains spectral components at the frequencies $pf_1 + qf_2$ where p and q are integers. Certain properties of these components can be derived for an arbitrary rectifier. Therefore, a specific analysis of the exponential rectifier is deferred for the moment. Let R[a(t)] denote a function relating an input signal to an output signal. This function is assumed to have no memory (i.e., the

function maps the amplitude maps the amplitude of its input to an output amplitude regardless of the time-behavior of the input). Let a(t), as given in equation 5.5, be the input with θ_{11} and θ_{21} set to zero. A period histogram synchronized to the frequency $pf_1 + qf_2$ is computed from the output of the rectifier. Goldstein (1972) has shown that the period histogram is expressed as

$$g_m(pf_1+qf_2) = \int_0^1 R\left[a_{11} \cos 2\pi q\alpha + a_{21} \cos 2\pi\left(p\alpha - \frac{m}{Mq}\right)\right] d\alpha \tag{5.6}$$

where

$g_m(pf_1+qf_2)$ represents the contents of the m^{th} bin, $m = 0, \ldots, M-1$, of a period histogram synchronized to the frequency $pf_1 + qf_2$.

The expression for the period histogram synchronized to $pf_1 - qf_2$ is related to equation (5.6) by the following

$$g_m(pf_1-qf_2) = \int_0^1 R\left[a_{11} \cos 2\pi q\alpha + a_{21} \cos 2\pi\left(p\alpha + \frac{m}{Mq}\right)\right] d\alpha$$

$$= \int_0^1 R\left[a_{11} \cos 2\pi q\alpha + a_{21} \cos 2\pi\left(p\alpha - \frac{(-m)}{Mq}\right)\right] d\alpha$$

$$g_m(pf_1-qf_2) = g_{-m}(pf_1+qf_2) \tag{5.7}$$

These period histograms are identical except for a reversal of the bin indicator m. Consequently, the computation of the synchronization indices from these period histograms yields

$$S_{pf_1+qf_2} = S_{pf_1-qf_2}, \qquad p, q \geq 1 \tag{5.8}$$

Components in the output of the rectifier at the frequencies nf_1+mf_2 and nf_1-mf_2 have the same amplitudes no matter what the form of the rectifier so long as it is memoryless.

Returning to the exponential rectifier, its output r(t) is written

$$r(t) = \exp\left\{a_{11}\cos 2\pi(f_1t+\theta_{11}) + a_{21}\cos 2\pi(f_2t+\theta_{21})\right\} \tag{5.9}$$

Using the relationship (Abramowitz and Stegun, 1965: 376)

$$\exp\{a\cos\alpha\} = I_o(a) + 2\sum_{k=1}^{\infty} I_k(a)\cos k\alpha, \tag{5.10}$$

where

$I_k(a)$ is the k^{th} order, modified Bessel function of the first kind.

Equation 5.9 is written

$$r(t) = I_o(a_{11})\, I_o(a_{21}) \cdot \Bigg\{ 1 + 2 \sum_{p=1}^{\infty} \frac{I_p(a_{11})}{I_o(a_{11})} \cos 2\pi p(f_1 t + \theta_{11}) + 2 \sum_{q=1}^{\infty} \frac{I_q(a_{21})}{I_o(a_{21})} \cos 2\pi q(f_2 t + \theta_{21}) + 2 \sum_{p=1}^{\infty} \sum_{q=1}^{\infty} \frac{I_p(a_{11})\, I_q(a_{21})}{I_o(a_{11})\, I_o(a_{21})} \Big[\cos 2\pi[(pf_1 + qf_2)t + p\theta_{11} + q\theta_{21}] + \cos 2\pi[(pf_1 - qf_2)t + p\theta_{11} - q\theta_{21}] \Big] \Bigg\} \tag{5.11}$$

The synchronization indices are given by the amplitudes of each component within the braces. Therefore

$$S_{pf_1 + qf_2} = \frac{I_p(a_{11})\, I_q(a_{21})}{I_o(a_{11})\, I_o(a_{21})} = S_{pf_1 - qf_2} \tag{5.12a}$$

$$\phi_{pf_1 \pm qf_2} = p\theta_{11} \pm q\theta_{21} \tag{5.12b}$$

where

$$p,\ q \geq 1$$

As $S_{pf_1} = \frac{I_p(a_{11})}{I_o(a_{11})}$, $\phi_{pf_1} = p\theta_{11}$, $S_{qf_2} = I_q(a_{21})/I_o(a_{21})$, and $\phi_{qf_2} = q\theta_{21}$, we find that for the exponential rectifier

$$S_{pf_1 \pm qf_2} = S_{pf_1} \cdot S_{qf_2} \qquad p, q \geq 1 \qquad (5.13a)$$

$$\phi_{pf_1 \pm qf_2} = \phi_{pf_1} \pm \phi_{qf_2} \qquad p, q \geq 1 \qquad (5.13b)$$

where these synchronization indices and phases are measures of the response to the two-tone stimulus. This set of equation are necessary and sufficient conditions for an exponential rectifier having an input consisting of two inharmonic sinusoids.

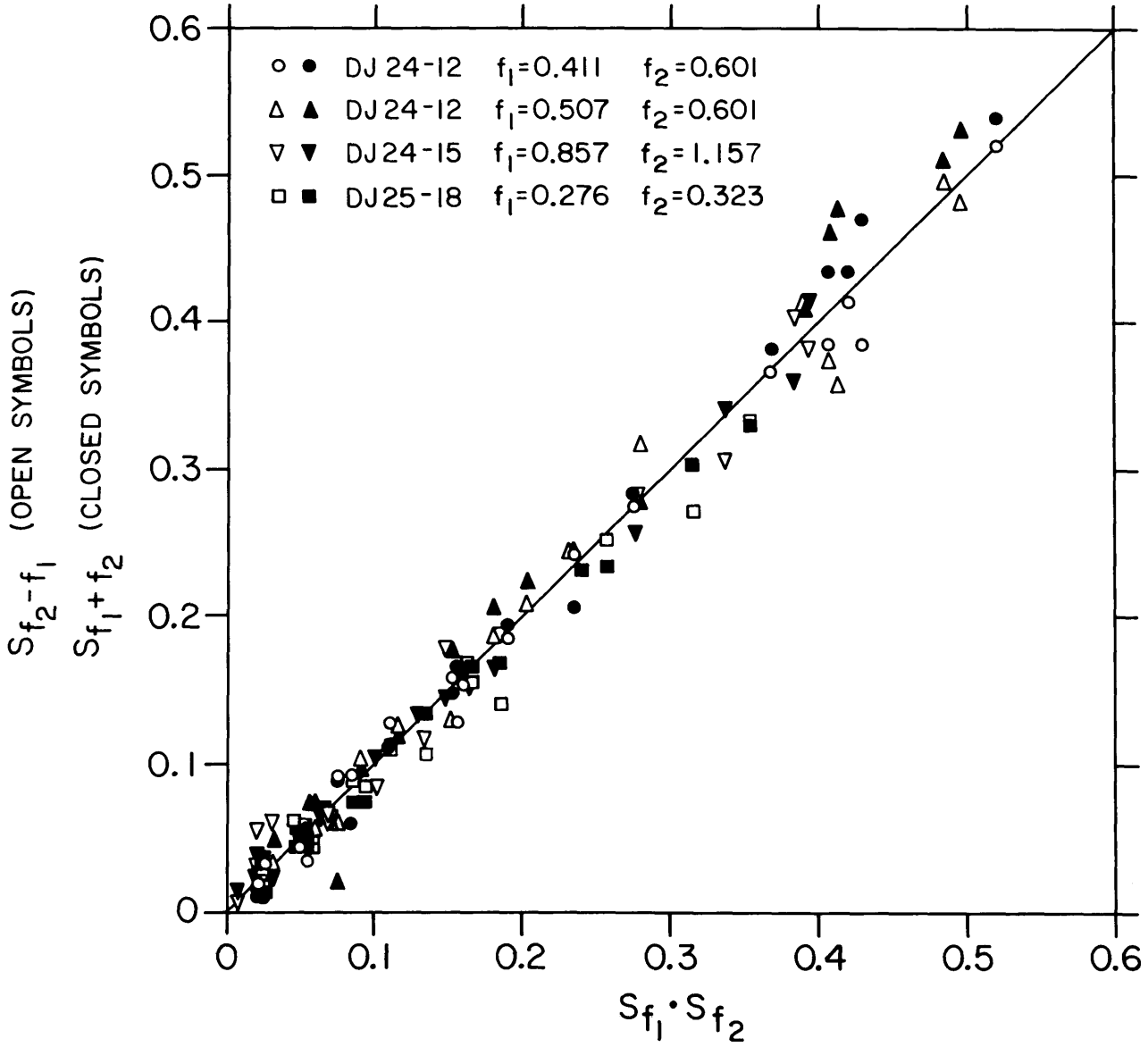
In section 5.3, the amplitude of the synchronized responses to the sum and difference frequencies were shown to be equal for low-level stimuli (Figure 5.5). Similar results have been found by Pfeiffer et al. (1974). These results are consistent with the assumption that the source of these components is the memoryless rectification of two sinusoids (equation 5.8). Any model which follows the rectification process by a system having memory (over the range of frequencies spanned by the data, i.e., frequencies less than 2 kHz) cannot be consistent with these data (e.g., Anderson's (1973) model is of this form. See section 6.1.1).

In order to be consistent with an exponential rectifier, equations 5.13a and 5.13b should describe the data. Figure 5.7 plots

Figure 5.7. Relationship between the Amplitudes of Frequency Distortion Components in the Synchronization Response and the Amplitudes of Components Synchronized to the Stimulus Frequencies

Response patterns were recorded from single auditory-nerve fibers responding to a two-tone stimulus. Period histograms synchronized to the stimulus frequencies (f_1 and f_2) and to the sum ($f_1 + f_2$) and difference ($f_2 - f_1$) were computed from these response patterns. Synchronization indices were then calculated from these period histograms.

The synchronization indices calculated with respect to $f_1 + f_2$ and $f_2 - f_1$ are plotted against the product of the synchronization indices to f_1 and f_2. The sum- and difference-frequency synchronization indices depicted here are the same as those displayed in Figure 5.5. The line denotes the equation $S_{f_2 \pm f_1} = S_{f_1} \cdot S_{f_2}$.


DJ 24-12 f_1=0.411 f_2=0.601
DJ 24-12 f_1=0.507 f_2=0.601
DJ 24-15 f_1=0.857 f_2=1.157
DJ 25-18 f_1=0.276 f_2=0.323
$S_{f_2-f_1}$ (OPEN SYMBOLS)
$S_{f_1+f_2}$ (CLOSED SYMBOLS)
$S_{f_1} \cdot S_{f_2}$
0
0.1
0.2
0.3
0.4
0.5
0.6

the synchronization indices for the sum and difference frequencies against the product of the synchronization indices measured at the stimulus frequencies. These data are consistent with equation 5.13a. Pfeiffer et al. have measured the phases as well as the synchronization indices to the sum and difference frequencies. They noted that their phase data are well described by equation 5.13b. These two sets of data are consistent with the assumption that these frequency distortion components are attributable to the exponential rectification of two sinusoids.

There are examples of synchronized responses to frequency distortion components that are not consistent with memoryless rectifier models. The exact nature of the inconsistentencies can be categorized in two ways. If equation 5.8 is not satisfied, the mechanism responsible for the distortion component cannot be described by the memoryless rectification of two sinusoids. If equation 5.13a is not satisfied, an exponential-rectifier model cannot describe the data. Obviously, if the first criterion is not satisfied (equation 5.8) then the second (equation 5.13a) will not be.

The data illustrated in Figure 3.8 provide an example of synchronized responses to the sum and difference frequencies which cannot be described by memoryless rectification. The synchronized response to the difference frequency is the largest component in the response. This response pattern was obtained at higher levels than those illustrated in Figure 5.5. If the levels of the two stimulus tones are not large, then the responses synchronized to the sum and difference frequencies can be described by the exponential rectifier

in the model depicted by Figure 4.22. At larger levels, the synchronized responses to the sum and difference frequencies cannot be described by the exponential rectification of two sinusoids. Earlier stages of the model could describe these data; in particular, the square-law device produces components at the sum and difference frequencies (equation 5.2). Whether or not the model described these data was not empirically tested.

The synchronized response to the combination frequency $2f_1-f_2$ can also be analyzed in this manner. If the responses synchronized to this frequency could be ascribed to exponential rectification, then

$$S_{2f_1-f_2} = S_{2f_1} \cdot S_{f_2} \tag{5.14a}$$

$$\phi_{2f_1-f_2} = 2\phi_{f_1} - \phi_{f_2} \tag{5.14b}$$

This equation implies that the combination-frequency response would be small in amplitude, smaller than the sum or difference frequency components.* No measurements of this response were made as a part of this thesis. Pfeiffer et al. (1974) report measurements of a response synchronized to $2f_1-f_2$. The amplitude of this component was less than either the sum or difference frequency

*For exponential rectification, $S_{nf_1} < S_{f_1}$, $n \geq 2$. Therefore $S_{2f_1} \cdot S_{f_2} < S_{f_1} \cdot S_{f_2}$ which implies $S_{2f_1 \pm f_2} < S_{f_1 \pm f_2}$.

components. Furthermore, they indicated that equation 5.14b was satisfied by their data. As far as these data have been analyzed, this synchronized response to $2f_1-f_2$ can be described by the exponential rectifcation of two sinusoids.

Just as with the difference-frequency responses, synchronized responses to $2f_1-f_2$ have been reported which cannot be attributable to an exponential rectifier.* Goldstein and Kiang (1968) report responses synchronized to $2f_1-f_2$ in auditory-nerve fibers responding to a two-tone stimulus. The stimulus frequencies f_1 and f_2 were chosen so that $2f_1-f_2$ coincided with the CF of the fiber. The resulting response synchronized to $2f_1-f_2$ was larger than the responses synchronized to either f_1 or f_2. Furthermore, this response can be cancelled by the addition to the stimulus of a tone having frequency $2f_1-f_2$. The amplitude and phase settings of this tone which were required for cancellation in one unit suffice for cancellation in other units having different CFs (Goldstein, 1970). These response characteristics are not consistent with memoryless rectification. In terms of the single-tone model, a $2f_1-f_2$ component dependent upon the stimulus tones would have to added to the input of the bandpass filter in order to describe these data. It is for this reason that this response at the frequency $2f_1-f_2$ is sometimes termed "stimulus-like."

---

*A similar analysis was performed by Goldstein (1972).

5.6 Summary

Two aspects of the synchronized response of an auditory-nerve fiber to two tones were considered. First, two-tone synchrony suppression was demonstrated. Unlike two-tone rate suppression, synchrony suppression can occur between any pair of tones so long as each evokes a synchronized response. The level of tone required to suppress the response to another tone is roughly constant for low to moderate levels of the suppressed tone.

Secondly, synchronized responses were shown to be present at frequencies other than those harmonic with the stimulus frequencies. The sum and difference frequency components were emphasized. The response amplitudes at these frequencies were shown to be roughly equal, even for rather large values of the sum frequency (2 kHz). Furthermore, these amplitudes depended upon the response amplitudes at the stimulus frequencies rather than the stimulus amplitudes.

These two aspects of the two-tone response were also found in the output of the model. Two-tone synchrony suppression was correlated with the saturation of the relationship between the synchronized response to one tone and the level of the tone. Components at the sum and difference frequencies in the two-tone response pattern were described by the exponential rectification of two sinusoids.

CHAPTER VI

DISCUSSION AND CONCLUSIONS

6.1 Characteristics of the Single-Tone Response Pattern

6.1.1 Frequency range of the synchronized response

Information concerning the acoustic stimulus can be conveyed to the central nervous system by temporal variations synchronized to the stimulus in the rates of discharge of auditory-nerve fibers. The synchronization index S_f, defined as the amplitude of the fundamental component of the response, is a measure of the synchronized response to a tone. The level-maximum of the synchronization index depends on the tone frequency and can be used to assess the frequency range over which a synchronized response occurs. Our measurements show that the level-maximum of S_f decreases as frequency increases (Figure 4.4). This dependence can be described roughly by $S_f = 1 - f/6$ in the range of tone frequencies (f) lying between 1 and 5 kHz. Above 5 kHz, negligible synchronized responses are observed.

It is important to know whether this observed decrease and the difficulty of measuring any synchronized responses at higher frequencies might be due to our measurement techniques. Our analysis of period histograms (section 3.5) show that the procedures employed in the computation of a period histogram are sufficient to allow detection of stimulus-locked responses for frequencies

well above 5 kHz. We conclude that the limitation on the frequency range over which a synchronized response can be detected is due to cochlear processes rather than to measurement errors.

Synchronization indices computed from responses obtained at large stimulus levels are presented in a recent paper by Anderson (1973). His data were recorded from single auditory-nerve fibers in the squirrel monkey. The synchronization index was computed from single-tone response patterns obtained at a fixed stimulus level of 90 dB SPL (at the tympanic membrane). His measurements were not explicitly stated to be the maximum values of the synchronization index at a particular frequency. There is some discrepancy between his data and the data presented in Figure 4.4; his measurements of the synchronization index are consistently smaller. For instance, his measurements performed at approximately 1 kHz lie between about 0.6 and 0.75 (Anderson, 1973: Figure 4); our measurements lie between 0.72 and 0.86 near 1 kHz (Figure 4.4). No plausible explanation for this discrepancy can be offered here.

Anderson (1973) proposed that the decrease in the synchronized response at higher frequencies could be explained by jitter in the propagation time of a spike in an auditory-nerve fiber; our data obtained from two-tone responses are not easily understood with this hypothesis. According to Anderson's hypothesis, rates of discharge measured at the recording electrode are related to the rate of discharge at the site of spike initiation by a linear, time-invariant, lowpass filter Assuming that maximum values of

synchronization indices were independent of frequency at the spike initiation site, the relationship of the measured maximum values of S_f to the stimulus frequency would be equal to the frequency response of the lowpass filter. Results obtained from two-tone response patterns show that the synchronization indices computed from period histograms synchronized to the sum and difference frequencies are approximately equal in amplitude even though these two frequencies differ widely. In particular, the synchronization index to the sum frequency of over 2 kHz was found to be approximately equal to the index for a difference frequency of 0.3 kHz. In terms of Anderson's model, the synchronization index to 2 kHz should be at least 20% smaller than the synchronization index to 0.3 kHz. Therefore, at the site of spike initiation the synchronization index to the sum frequency would have to be 20% larger than the synchronization index to the difference frequency in order for his model to be consistent with our data. While these arguments do not disprove his hypothesis, it is more parsimonious to assume that the decline of synchronized responses with frequency occurs in processes peripheral to or at the point of spike initiation. Measurements of synchronized responses from, for instance, the cell bodies of the spiral ganglion would directly test his hypothesis.

6.1.2 Relationships of the synchronized response to the average-rate response

The data depicted in Figure 4.3 indicate that the average rate and the synchronization index are monotonically related to stimulus

level only over a restricted range of levels, some 20-30 dB, for levels less than 80 dB SPL. The average-rate data of Kiang (Kiang et al., 1965: 79-83; Kiang and Moxon, 1972) and Rose (Rose et al., 1971) and the synchronized response data of Littlefield (Littlefield et al., 1972) and Hind (Hind et al., 1967) support this view. Some investigators have reported somewhat larger ranges of level over which the average rate increases (Sachs et al., 1974; Geisler et al., 1974). In our data, there were instances when the maximum value of the average discharge rate could not be well defined. Whether this lack of definition was due to statistical variations in the measurement of the average rate or to the relationship of the average rate to the level of the stimulus could not be discerned.

A synchronized response can be measured at stimulus levels 10 to 20 dB below levels evoking a detectable increase in the average rate above the spontaneous rate (Figure 4.3). Similar results have been reported by other investigators (Littlefield et al., 1972; Rose et al., 1971). As the stimulus level is increased, both the average discharge rate and the synchronization index increases. Once the synchronization index attains its maximum value, the average discharge rate may increase over another 10 to 20 dB range of levels before attaining its maximum value (Figure 4.3). Hence, our data show that the range of levels over which a change in the response pattern (either the synchronized response or the average discharge rate) can be discerned is about 30-50 dB.

This result implies that at least a portion of the average rate

response at low frequencies is not related responses synchronized with the stimulus. At high frequencies, where no synchronized response is measured, the average discharge rate is entirely unrelated to a synchronized response. In the frequency region where the level-maximum of the synchronized response decreases (frequencies greater than 1 kHz), the average rate increases independently of the waveform of the synchronized response; these increases are accomplished by the multiplication of the waveform by a constant (Figure 4.7).

In terms of our model, these effects were lumped into a single, time-independent parameter having amplitude a_o. As this parameter occurs as an additive term in the argument of the final exponential, variation of a_o results in changes in the amplitude of the response waveform. This parameter varied with the level of the stimulus no matter what the frequency of the stimulus. Rather than describing the response to a tone in a disjoint manner (synchronized response at low frequencies, average rate response at high frequencies), it is more appropriate to emphasize the continuity of the response over the audio-frequency range. The synchronized response becomes less pronounced as the stimulus frequency is increased; throughout the entire range of frequencies, the average rate of discharge can change without concomitant changes in the synchronized response.

Since the average rate and the synchronized response can vary independently for a single-tone stimulus, it is conceivable that similar characteristics are found in the responses to other stimuli.

For instance, two-tone response patterns to low-frequency tones show two "inhibitory" phenomena: two-tone synchrony suppression and two-tone rate suppression. It may well be that these phenomena are completely unrelated. This view is examined more closely in the next section.

6.2 Characteristics of Two-Tone Response Patterns

6.2.1 Two-tone synchrony suppression and its relationship to single-tone responses

The suppression of the synchronized response to one tone by the addition of a second tone to stimulus is found in auditory-nerve fiber response patterns and in the output of the model (Figures 5.3 and 5.6). In terms of the model, the sources of two-tone synchrony suppression are the automatic gain controls. These elements are required in the model so as to reflect the saturation of the synchronized response to one tone as a function of level (Figure 4.3). Therefore, two-tone synchrony suppression and the saturation with level of the synchronized response are related phenomena as they are traceable to the same elements in the model.

An analysis of two-tone synchrony suppression has also been performed by Rose et al. (1974). They considered how the synchronized response to a single tone deviated from a linear relationship with stimulus level. Using this comparison for each tone in a two-tone stimulus, they inferred the degree of two-tone synchrony suppression. Their result is similar to the one presented here: two-tone synchrony suppression can be predicted from considerations of how the synchronized response saturates with level.

6.2.2 Two-tone synchrony suppression and its possible relation to two-tone rate suppression

It was suggested in section 6.1.2 that two-tone synchrony suppression and two-tone rate suppression may be unrelated phenomena. As we do not have reliable average-rate data from our experiments involving two tones, we analyze in this section the relationships between these phenomena in two ways: 1) by a comparison of data obtained by different investigators and 2) by a formulation of how two-tone rate suppression would have to be incorporated into the single-tone model.

Two-tone rate suppression depends upon the relationship of the stimulus frequencies to the CF of the fiber (Sachs and Kiang, 1968). Given one tone at the CF, the second tone must lie in the suppression areas flanking the slopes of the tuning curve. Two-tone synchrony suppression does not depend upon the choice of the stimulus frequencies relative to the CF (Hind et al., 1967) so long as each tone evokes a synchronized response. Two-tone rate suppression occurs for units whose CFs cover a wide frequency range and the characteristics of this phenomenon do not vary, at least qualitatively speaking, upon the absolute values of the frequencies involved (Sachs and Kiang, 1968; Sachs, 1969; Hind, 1970). Two-tone synchrony suppression occurs whenever each tone evokes a synchronized response (i.e., for tone frequencies less than about 5 kHz).

On the basis of these empirical results, the following deductions can be drawn. If two-tone synchrony suppression were involved in two-tone rate suppression, one might expect the characteristics of the rate suppression phenomenon to vary with frequency. Also, the characteristics of two-tone synchrony suppression would need to vary with the relationships of the stimulus frequencies to the rate-suppression areas of the fiber. Neither of these stimulus-response relationships have been reported and neither has been observed in the course of this study.

This apparent lack of correspondence between these response phenomena can be incorporated into the model. Two-tone synchrony suppression is described by the stimulus-locked branches of the model. In order to describe two-tone rate suppression in the framework of the model, the time-independent term a_o would have to reflect the entire phenomenon. By allowing this term to vary not only with level but also with the precise relationships of the stimulus frequencies to the CF of the fiber, it is conceivable that the model can encompass two-tone rate suppression. The resulting variation of a_o would not influence the stimulus-locked branches of the model. Therefore, whether two-tone rate suppression occurs or not, the synchronized response would not be affected.

To the accuracy of the measurements and to the degree the model is capable of describing two-tone responses, two-tone rate suppression and two-tone synchrony suppression are unrelated phenomena.

6.2.3 Frequency Distortion Components

Using the model, we have related responses synchronized to the sum and difference frequencies to the waveform of the response to a single tone. In the model, these frequency distortion components are due to the exponential rectification of two sinusoids. A rectifier is required in the model to describe the waveform of the response to one tone. In the case of single tone responses, harmonics of the stimulus tone are found in the response (Figure 4.2). These harmonics are a consequence of the shape of the waveform and are primarily due to the rectifier. When the stimulus consists of two tones having frequencies f_1 and f_2, the correlate of the harmonics in the single-tone responses is the generation of frequency components in the response at frequencies $pf_1 + qf_2$. (See equation 5.12a). As with two-tone synchrony suppression, we find an example of a two-tone response phenomenon which can be derived from considerations of single-tone responses patterns.

However, there are responses synchronized to $pf_1 + qf_2$ which cannot be described by the simple rectification of two sinusoids. Certain difference-frequency responses (Figure 3.8) are inconsistent with this notion; however, it is conceivable that the square-law device in the model could account for this response. There was insufficient data to test this hypothesis. A more serious exception to this simple model are responses synchronized to the combination-frequency $2f_1 - f_2$ (section 5.5.2). Presumably, this

response can originate in earlier stages of the model while retaining the rectifier as the final element. These results suggest that there are two categories of distortion components – those generated by the exponential rectification of two sinusoids and those components that already exist at the input to the exponential rectifier. This latter category of response is present when higher stimulus levels are used and the stimulus frequencies lie above the CF of the fiber.

6.3 General Remarks

Based upon a systematic study of the responses to a single tone, we formulated a model to describe the details of these responses. The resulting model, based upon the exponential approximation (equation 4.4), describes well the single-tone response. The framework of the analytical model, as depicted in Figure 4.22, is not intended to refer to sequence of events in the peripheral auditory system which leads from a sinusoidal sound pressure at the tympanic membrane to discharge patterns in auditory-nerve fibers. The events comprising this transduction process are presently not well understood; these processes are subjects of active research. The physiological mechanisms of the cochlea cannot be inferred solely from data obtained from auditory-nerve fibers. Such an inference is equivalent to a specification of the elements of a system from its input/output relationship. In general, that sort of analysis does not yield a unique result. In the case of cochlear physiology, we are ultimately interested in cochlear mechanisms, not in an equivalent system.

Rather, the model developed in this study has uses in other areas. First of all, it aids in understanding the components of the single-tone response. For instance, this response is comprised of two basic components: one which is not synchronized with the stimulus and one which is. The component synchronized to the stimulus becomes less effective as the tone frequency is increased while the time-independent component is present at all frequencies. The decrease in the synchronized component with increasing frequency is

related to changes in the skew and asymmetry of the waveform of the rate of discharge. As the skew is due to the second harmonic and as the asymmetry is primarily due to the exponential rectification of the fundamental, their decreases with increasing frequency are traced to the same element of the model: the lowpass filter $H_L(f)$ (Figure 4.22). This filter accounts for the lack of synchronized responses at high frequencies.

Secondly, the model is used to determine whether the responses to other stimuli can be deduced from the responses to single tones. If they can be deduced, then these responses can be placed into a common theoretical framework. In this study, the responses to two tones are analyzed with respect to the model. Some aspect of these responses (two-tone synchrony suppression and the generation of certain frequency distortion components) are predicted by the model while other aspects (e.g., combination-frequency $(2f_1-f_2)$ responses as reported by Goldstein and Kiang (1968)) are not. The fact that the responses to stimuli other than single tones are consistent with predictions of the model lends credence to the model.

Those aspects predicted by the model do not provide more information about the stimulus-response relationships of the peripheral auditory system than do single-tone response patterns. It is those responses not predicted by the model which reveal more about the general properties of the responses of auditory-nerve fibers to acoustic stimuli. A study of the responses not predicted by the model is the logical course of research in view of the results presented here.

REFERENCES

Abramowitz, M., and Stegun, I.A. (1965). Handbook of Mathematical Functions (Dover Publications, New York).

Anderson, D.J. (1973). "Quantitative model for the effects of stimulus frequency upon synchronization of auditory nerve discharges," J. Acoust. Soc. Am. 54: 361-364.

Anderson, D.J., Rose, J.E., Hind, J.E., and Brugge, J.F. (1971). "Temporal position of discharges in single auditory nerve fibers with the cycle of a single-wave stimulus: frequency and intensity effects," J. Acoust. Soc. Am. 49: 1131-1139.

Borg, E. (1972). "Acoustic middle ear reflexes: a sensory-control system," Acta Oto-laryng., Suppl. 304.

Carmel, P.W., and Starr, A. (1963). "Acoustic and non-acoustic factors modifying middle-ear muscle activity in waking cats," J. Neurophysiol. 26: 598-616.

Citron, L., Exley, D., and Hallpike, C.S. (1956). "Formation, circulation and chemical properties of the labyrinthine fluids," Brit. Med. Bull. 12: 101-104.

Colburn, H.S. (1973). "Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination," J. Acoust. Soc. Am. 54: 1458-1470.

Cox, D.R. (1962). Renewal Theory, in Methuen's Monographs on Applied Probability and Statistics (John Wiley and Sons, Inc., New York).

Cox, D.R., and Lewis, P.A.W. (1966). The Statistical Analysis of Series of Events, in Methuen's Monographs on Applied Probability and Statistics (John Wiley and Sons, Inc., New York).

Cramér, H. (1946). Mathematical Methods of Statistics (Princeton University Press, Princeton, N.J.).

Dallos, P. (1970). "Low-frequency auditory characteristics: species dependence," J. Acoust. Soc. Am. 48: 489-499.

Dallos, P. (1973). The Auditory Periphery: Biophysics and Physiology (Academic Press, New York).

Davis, H. (1958). "A mechano-electrical theory of cochlear action," Ann. Otol. Rhinol Laryng. 67: 789-801.

Davis, H. (1959). "Excitation of auditory receptors," Chapter XXIII in Handbook of Physiology, Section 1: Neurophysiology. Vol. I (Waverly Press, Inc., Baltimore): 565-584.

Evans, E.F. (1972). "The frequency response and other properties of single fibres in the guinea-pig cochlear nerve," J. Physiol. 226: 263-287.

Gacek, R.R., and Rasmussen, G.L. (1961). "Fiber analysis of the statoacoustic nerve of guinea pig, cat and monkey," Anat. Rec. 139: 455-463.

Geisler, C.D. (1968). "A model of the peripheral auditory system responding to low-frequency tones," Biophys. J. 8. 1-15.

Geisler, C.D., Rhode, W.S., and Kennedy, D.T. (1974). "The responses to tonal stimuli of single auditory nerve fibers and their relationship to basilar membrane motion in the squirrel monkey," submitted to J. Neurophysiol.

Goldberg, J.M., and Brown, P.R. (1969). "Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization," J. Neurophysiol. 32: 613-636.

Goldstein, J.L. (1970). "Aural combination tones," in Frequency Analysis and Periodicity Detection in Hearing, edited by R. Plomp and G.F. Smoorenburg (A.W. Sijthoff, Leiden): 230-245.

Goldstein, J.L. (1972). "Neural phase locking to combination tones predicted by simple transducer models," J. Acoust. Soc. Am. 52: 142(A).

Goldstein, J.L., and Kiang, N. Y.-S. (1968). "Neural correlates of the aural combination tone $2f_1-f_2$," Proc. IEEE 56: 981-992.

Gray, P.R. (1966). "A statistical analysis of electrophysiological data from auditory nerve fibers in cat," M.I.T. R.L.E. Technical Report No. 451, Cambridge, Mass.

Greenwood, J.A., and Durand, D. (1955). "The distribution of length and components of the sum of n random unit vectors," Ann. Math. Stat. 26: 233-246.

Guinan, J.J., Jr., and Peake, W.T. (1967). "Middle-ear characteristics of anesthetized cats," J. Acoust. Soc. Am. 41: 1237-1261.

Hildebrand, F.B. (1956). Introduction to Numerical Analysis (McGraw-Hill Book Co., New York).

Hind, J.E. (1970). "Two-tone marking effects in squirrel monkey auditory nerve fibers," in Frequency Analysis and Periodicity Detection in Hearing, edited by R. Plomp and G.F.Smoorenburg (A.W. Sijthoff, Leiden): 193-200.

Hind, J.E., Anderson, D. J., Brugge, J.F., and Rose, J.E. (1967). "Coding of information pertaining to paired low frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol. 30: 794-816.

Johnson, D.H. (1970). "Statistical relationships between firing patterns of two auditory-nerve fibers," S.M. Thesis, Dept. of Electrical Engineering, M.I.T., Cambridge, Mass.

Johnstone, B.M., Taylor, K.J., and Boyle, A.J. (1970). "Mechanics of the guinea pig cochlea," J. Acoust. Soc. Am. 47: 504-509.

Kiang, N.Y.-S. (1968). "A survey of recent developments in the study of auditory physiology," Ann. Otol. Rhinol. Laryng. 77: 656-675.

Kiang, N.Y.-S. (1974). Personal communication.

Kiang, N.Y.-S., Baer, T., Marr, E.M., and Demont, D. (1969). "Discharge patterns of single auditory-nerve fibers as functions of tone level," J. Acoust. Soc. Am. 46: 106(A).

Kiang, N.Y.-S., and Moxon, E.C. (1972). "Physiological considerations in artificial stimulation of the inner ear," Ann. Otol. Rhinol. Laryng. 81: 714-730.

Kiang, N.Y.-S., Moxon, E.C., and Levine, R.A. (1970). "Auditory-nerve activity in cats with normal and abnormal cochlea," in Ciba Foundation Symposium on Sensorineural Hearing Loss, edited by G.E.W. Wolstenholme and J. Knight (J. & A. Churchill, London): 241-273.

Kiang, N.Y.-S., Watanabe, T., Thomas, E.C., and Clark, L.F. (1965). Discharge Patterns of Single Fibers in the Cat's Auditory Nerve, Research Monograph No. 35 (The M.I.T. Press, Cambridge, Mass.).

Kim, D.O., Littlefield, W.M., Molnar, C.E., and Pfeiffer, R.R. (1973a). "A model for discharge activity of single cochlear nerve fibers," J. Acoust. Soc. Am. 54: 283(A).

Kim, D.O., Molnar, C.E., and Pfeiffer, R.R. (1973b). "A system of nonlinear differential equations modeling basilar membrane motion," J. Acoust. Soc. Am. 54: 1517-1529.

Kimura, R. (1966). "Hairs of the cochlear sensory cells and their attachment to the tectorial membrane," Acta Oto-Laryng. 61: 55-72.

Kohllöffel, L.U.E. (1972). "A study of basilar membrane vibrations. III. The basilar membrane frequency response curve in the living guinea pig," Acustica 27: 82-89.

Littlefield, W.M., Pfeiffer, R.R., and Molnar, C.E. (1972). "Modulation index as a response criterion for discharge activity," J. Acoust. Soc. Am. 51: 93(A).

Lorente de Nó, R. (1937). "The sensory endings in the cochlea," Laryngoscope 47: 373-377.

Møller, A.R. (1962). "The sensitivity of contraction of the tympanic muscles in man," Ann. Otol. Rhinol. Laryng. 71: 86-95.

Moxon, E.C. (1967). "Electric stimulation of the cat's cochlea: a study of discharge rates in single auditory nerve fibers," S.M. Thesis, Dept. of Electrical Engineering, M.I.T., Cambridge, Mass.

Nedzelnitsky, V. (1974). "Measurements of sound pressure in the cochleae of anesthetized cats," Ph.D. Thesis, Dept. of Electrical Engineering, M.I.T., Cambridge, Mass.

Oppenheim, A.V., and Schafer, R.W. (1975). Digital Signal Processing (Prentice-Hall, Inc., Englewood Cliffs, N.J.).

Perkins, R.E. (1973). "Innervation patterns in cochleas of cat and rat: study with rapid Golgi techniques," Anat. Rec. 175: 410(A).

Pfeiffer, R.R., and Kim, D.O. (1973). "Considerations of nonlinear response properties of single cochlear nerve fibers," in Basic Mechanisms in Hearing, edited by A. Møller (Academic Press, New York): 555-587.

Pfeiffer, R.R., and Molnar, C.E. (1970). "Cochlear nerve fiber discharge patterns: relationship to the cochlear microphonic," Science 167: 1614-1616.

Pfeiffer, R.R., Molnar, C.E., and Cox, J.R. (1974). "The representation of tones and combination tones in spike discharge patterns of single cochlear nerve fibers," in Facts and Models in Hearing, edited by E. Zwicker and E. Terhardt (Springer-Verlag, New York): 323-331.

Rasmussen, G.L. (1960). "Efferent fibers of the cochlear nerve and cochlear nucleus," Chapter 8 in Neural Mechanisms of the Auditory and Vestibular Systems, edited by G.L. Rasmussen and W.F. Windle (Charles C. Thomas, Springfield, Ill.): 105-115.

Rayleigh, J.W.S. (1905). "The problem of the random walk," Nature 72: 318.

Rhode, W.S. (1971). "Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique," J. Acoust. Soc. Am. 49: 1218-1231.

Rhode, W.S., and Robles, L. (1974). "Evidence from Mössbauer experiments for nonlinear vibration in the cochlea," J. Acoust. Soc. Am. 55: 588-596.

Rodieck, R.W., Kiang, N.Y.-S., and Gerstein, G.L. (1962). "Some quantitative methods for the study of spontaneous activity of single neurons," Biophys. J. 2: 351-368.

Rose, J.E. (1970). "Discharges of single fibers in the mammalian auditory nerve," in Frequency Analysis and Periodicity Detection in Hearing, edited by R. Plomp and G.F. Smoorenburg (A.W. Sijthoff, Leiden): 176-188.

Rose, J.E., Brugge, J.F., Anderson, D.J., and Hind, J.E. (1967). "Phase-locked response of low-frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol. 30: 769-793.

Rose, J.E., Hind, J.E., Anderson, D.J., and Brugge, J.F. (1971). "Some effects of stimulus intensity on response of auditory nerve fibers in the squirrel monkey," J. Neurophysiol. 34: 685-699.

Rose, J.E., Kitzes, L.M., Gibson, M.M., and Hind, J.E. (1974). "Observations on phase-sensitive neurons of the anteroventral cochlear nucleus of the cat: nonlinearity of cochlear outputs," J. Neurophysiol. 37: 218-253.

Sachs, M.B. (1969). "Stimulus-response relation for auditory-nerve fibers: two-tone stimulus," J. Acoust. Soc. Am. 45: 1025-1036.

Sachs, M.B., Abbas, P.J., and Bevan, M.F. (1974). "Discharge rate as a function of stimulus level for tones and noise in cat auditory-nerve fibers," J. Acoust. Soc. Am. 55: 466(A).

Sachs, M.B., and Kiang, N.Y.-S. (1968). "Two-tone inhibition in auditory-nerve fibers," J. Acoust. Soc. Am. 43: 1120-1128.

Schroeder, M.R., and Hall, J.L. (1974). "Model for mechanical to neural transduction in the auditory receptor," J. Acoust. Soc. Am. 55: 1055-1060.

Siebert, W.M. (1970). "Frequency discrimination in the auditory system: place or periodicity mechanisms?" Proc. IEEE 58: 723-730.

Simmons, F.B. (1959). "Middle ear muscle activity at moderate sound levels," Ann. Otol. Rhinol. Laryng. 68: 1126-1144.

Smith, C.A., Lowry, O.H., and Wu, M.-L. (1954). "The electrolysis of the labyrinthine fluids," Laryngoscope 64: 141-153.

Spoendlin, H. (1966). The Organization of the Cochlear Receptor in Advances in Oto-Rhino-Laryngology, Vol. 13 (S. Karger, New York).

Spoendlin, H. (1969). "Innervation patterns in the Organ of Corti of the cat," Acta Oto-Laryng. 67: 239-254.

Spoendlin, H. (1972). "Innervation densities of the cochlea," Acta Oto-Laryng. 72: 235-248.

Spoendlin, H., and Gacek, R.R. (1963). "Electromicroscopic study of the efferent and afferent innervation of the Organ of Corti in the cat," Ann. Otol. Rhinol. Laryng. 72: 660-686.

Tonndorf, J., and Khanna, S.M. (1968). "Submicroscopic displacement amplitudes of the tympanic membrane (cat) measured by laser interferometer," J. Acoust. Soc. Am. 44: 1546-1554.

Van Trees, H.L. (1968). Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory (John Wiley and Sons, Inc., New York).

von Békésy, G. (1960). Experiments in Hearing (McGraw-Hill Book Co., New York).

von Békésy, G., and Rosenblith, W.A. (1951). "The mechanical properties of the ear," in Handbook of Experimental Psychology, edited by S.S. Stevens (John Wiley and Sons, New York).

Weiss, T.F. (1966). "A model of the peripheral auditory system," Kybernetik 3: 153-175.

Wiener, F.M., Pfeiffer, R.R., and Backus, A.S.N. (1966). "On the sound pressure transformation by the head and auditory meatus of the cat," Acta Oto-Laryng. 61: 255-269.

Wiener, F.M., and Ross, D.A. (1946). "The pressure distribution in the auditory canal in a progressive sound field," J. Acoust. Soc. Am. 18: 401-408.

Wilson, J.P. (1974). "Basilar membrane vibration data and their relation to theories of frequency analysis," in Facts and Models in Hearing, edited by E. Zwicker and E. Terhardt (Springer-Verlag, New York): 56-63.

APPENDIX I

COMPUTATION OF THE EFFECT OF RANDOM BASELINE NOISE ON THE RATE OF DISCHARGE

If r(t), the rate of discharge of a spike train, is given by

$$r(t) = \bar{r}[1 + 2S_f \cos 2\pi ft] \tag{A1.1}$$

where

$\bar{r}$ is the average rate of discharge.

S_f is the amplitude of the synchronized component.

and if the baseline signal n(t) is a stationary, stochastic process with a Gaussian amplitude-density having variance σ_n^2 and zero mean, then the rate of discharge $\tilde{r}(t)$ of a train of pulses triggered from the recorded spike train is expressed by

$$\tilde{r}(t) = \bar{r} \cdot \left[1 + \frac{\dot{n}(t)}{m_s} \right] \left[1 + 2S_f \cos 2\pi f \left(t + \frac{n(t)}{m_s} \right) \right] \tag{A1.2}$$

The purpose of this appendix is the computation of the expected value of $\tilde{r}(t)$ with respect to n(t). Expanding equation (A1.2), we have

$$\tilde{r}(t) = \bar{r} \left[1 + \frac{\dot{n}(t)}{m_s} + 2S_f \cos 2\pi f \left(t + \frac{n(t)}{m_s} \right) + \frac{2S_f \dot{n}(t)}{m_s} \cos 2\pi f \left(t + \frac{n(t)}{m_s} \right) \right] \tag{A1.3}$$

Let us evaluate the expected value of (A1.3) term by term. As n(t) has zero mean

$$E[\dot{n}(t)] = 0 \tag{A1.4}$$

The expected value of the third term is written

$$E\left[\cos 2\pi f\left(t+\frac{n(t)}{m_s}\right)\right] = E\left[\cos 2\pi ft \cos\frac{2\pi fn(t)}{m_s} - \sin 2\pi ft \sin\frac{2\pi fn(t)}{m_s}\right] \tag{A1.5}$$

We note that the expected value of an odd function of n(t) is zero. Consequently, the second term of equation (A1.5) is zero. We are left with

$$E\left[\cos 2\pi f\left(t+\frac{n(t)}{m_s}\right)\right] = \cos 2\pi ftE\left[\cos\frac{2\pi fn(t)}{m_s}\right] \tag{A1.6}$$

This expected value is easily evaluated to yield

$$E\left[\cos 2\pi f\left(t+\frac{n(t)}{m_s}\right)\right] = \exp\left\{-\frac{1}{2}\left(\frac{2\pi f\sigma_n}{m_s}\right)^2\right\} \cos 2\pi ft \tag{A1.7}$$

Returning to the fourth term in equation (A1.3), we need only evaluate

$$E\left[\dot{n}(t) \cos 2\pi f\left(t+\frac{n(t)}{m_s}\right)\right] = \cos 2\pi ft \cdot E\left[\dot{n}(t) \cos\frac{2\pi fn(t)}{m_s}\right] - \sin 2\pi ftE\left[\dot{n}(t) \sin\frac{2\pi fn(t)}{m_s}\right] \tag{A1.8}$$

The first term in equation (A1.8) is zero because the term in brackets is an odd function of n(t). Therefore, we need only to evaluate the second term.

$$E\left[\dot{n}(t)\ \sin\frac{2\pi fn(t)}{m_s}\right] = \lim_{\Delta t\to 0} E\left[\frac{(n(t+\Delta t)-n(t))}{\Delta t}\cdot \sin\frac{2\pi fn(t)}{m_s}\right]$$

$$E\left[n(t)\ \sin\frac{2\pi fn(t)}{m_s}\right] = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{ E\left[n(t+\Delta t)\cdot \sin\frac{2\pi fn(t)}{m_s}\right] - E\left[n(t)\ \sin\frac{2\pi fn(t)}{m_s}\right]\right\} \qquad (A1.9)$$

The components $n(t+\Delta t)$ and $n(t)$ are jointly Gaussian random variables having correlation coefficient $R_n(\Delta t)/\sigma_n^2$ ($R_n(\tau)$ denotes the autocorrelation function of $n(t)$). Consequently,

$$E\left[n(t+\Delta t)\ \sin\frac{2\pi fn(t)}{m_s}\right] = \frac{R_n(\Delta t)}{\sigma_n^2}\cdot E\left[n(t)\ \sin\frac{2\pi fn(t)}{m_s}\right] \qquad (A1.10)$$

Therefore, equation (A1.9) can be written

$$E\left[n(t)\ \sin\frac{2\pi fn(t)}{m_s}\right] = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{\frac{R_n(\Delta t)}{\sigma_n^2} - 1\right\}\cdot E\left[n(t)\ \sin\frac{2\pi fn(t)}{m_s}\right]$$

$$E\left[n(t)\ \sin\frac{2\pi fn(t)}{m_s}\right] = \lim_{\Delta t\to 0}\frac{1}{\sigma_n^2}\left\{\frac{R_n(\Delta t)-\sigma_n^2}{\Delta t}\right\}\cdot E\left[n(t)\ \sin\frac{2\pi fn(t)}{m_s}\right] \qquad (A1.11)$$

As $R_n(0) = \sigma_n^2$, we see that the coefficient of the expected value in equation (A1.11) is the derivative of $R_n(\tau)$ at the origin.

$$E\left[n(t)\ \sin\frac{2\pi fn(t)}{m_s}\right] = \left.\frac{R_n'(\tau)}{\sigma_n^2}\right|_{\tau=0}\cdot E\left[n(t)\ \sin\frac{2\pi fn(t)}{m_s}\right] \qquad (A1.12)$$

If the derivative is defined, $R_n'(\tau)\big|_{\tau=0}$ is zero since $R_n(\tau)$ is an even function. Consequently, this term is zero.

Combining the results of equations (A1.4), (A1.7) and (A1.12), we can evaluate the expected value of $\tilde{r}(t)$ (equation (A1.3)).

$$E[\tilde{r}(t)] = \bar{r} \cdot \left[1 + 2S_f \exp\left\{ -\frac{1}{2}\left(\frac{2\pi f \sigma_n}{m_s}\right)^2 \right\} \cos 2\pi f t \right] \qquad (A1.13)$$

APPENDIX II

EVALUATION OF THE STATISTICS OF THE SYNCHRONIZATION INDEX AND THE PHASE

An analytical evaluation of the statistics of S_f and ϕ_f measured from any period histogram is a difficult task. However, some results are available. Based upon an asymptotic formula due to Rayleigh (1905), the expected value and variance of $\hat{S}_f$ for* the case $S_f = 0$ is given by

$$E[\hat{S}_f] = \left(\frac{\pi}{4N}\right)^{1/2} \tag{A2.1a}$$

$$\sigma^2_{\hat{S}_f} = \frac{4 - \pi}{4N} \tag{A2.1b}$$

where

N is the number of responses in the period histogram.

Equation A2.1a indicates the presence of a potentially significant bias in the measurement of S_f. For example, if N = 100, $E[\hat{S}_f] = 0.09$. In general, only values of $\hat{S}_f$ greater than about 0.1 can be considered relatively free of bias. Greenwood and Durand (1955) provide a formula for the pdf of $\hat{S}_f$ given that the waveform of the period histogram is described by

$$\lambda(t) = \lambda_o \exp\{a \sin 2\pi(ft+\phi_f)\} \tag{A2.2}$$

*For the purposes of this appendix, the measured value $\hat{S}_f$ of the synchronization index is distinguished from the actual value S_f. A similar distinction applies to $\hat{\phi}_f$ and ϕ_f.

However, an explicit computation of the statistics of $\hat{S}_f$ from their formula is difficult. Consequently, a simulation of the statistics of a period histogram was performed and estimates of the standard deviation of $\hat{S}_f$ and ϕ_f were then computed. Two basic components comprise a model of the statistics of a period histogram

1. a probabilistic description of the period histogram
2. a specification of the expected value of the histogram.

The probabilistic description used was originated by J. E. Evans (unpublished work) and W. M. Littlefield et al. (1972). Given a Poisson process description of the discharge pattern, the probability of a spike's time of occurrence being quantized to the m^{th} bin of a period histogram is independent of the bin in which another spike is placed and is given by equation 3.23.

$$\text{Prob}\,[x_m = 1] = \int_{m\Delta}^{(m+1)\Delta} \lambda(a)\, da \qquad \text{(A2.3)}$$

For simplicity, the binwidth Δ is assumed to be small relative to the variations in $\lambda(t)$. Consequently, equation A2.3 is approximated by

$$\text{Prob}\,[x_m = 1] \approx \lambda(m\Delta)\,\Delta \qquad \text{(A2.4)}$$

The expected value and variance of g_m, the contents of the m^{th} bin in the period histogram, is given by

$$E[g_m] = R\lambda(m\Delta)\,\Delta \qquad \text{(A2.5a)}$$

$$\text{var}[h_m] = R\lambda(m\Delta)\,\Delta \cdot [1 - \lambda(m\Delta)\Delta] \qquad \text{(A2.5b)}$$

where

R is the number of stimulus markers

As R is of the order of 10^4 and as the number of spikes in a bin of a histogram computed from auditory-nerve fiber discharge pattern is less than 100,* the numerical value of $\lambda(m\Delta)$ is usually less than 0.01. To a good approximation, the variance of g_m may be written

$$\text{var}\,[g_m] \approx R \cdot \lambda(m\Delta)\,\Delta \tag{A2.6}$$

Therefore

$$\text{var}\,[g_m] \approx E[g_m] \tag{A2.7}$$

The rate of discharge of an auditory-nerve unit is assumed to be given by equation A2.2

$$\lambda(t) = \lambda_o \exp\{a \sin 2\pi(ft+\phi_f)\} \tag{A2.2}$$

This description is the same as proposed by Siebert (1970). While this description does not mimic the details of measurements of the rate of discharge, the gross features of the rate of discharge are approximated (Colburn, 1973). The synchronization index of $\lambda(t)$ is given by (Abramowitz and Stegun, 1965: 374-376):

$$S_f = \frac{I_1(a)}{I_o(a)} \tag{A2.8}$$

where

$I_n(a)$ is the n^{th} order modified Bessel fuction of the first kind with the argument equal to a.

The simulation computed period histograms according to the equation

*The observation interval is assumed to be 30 seconds or less and the number of bins in the period histogram is assumed to be large (~100).

$$g_m = \mathrm{RND}\left[s_m + w_m\sqrt{s_m}\right] \qquad m = 0, \ldots, M-1 \qquad \text{(A2.9)}$$

where

$$s_m = R\lambda_o\Delta \exp\left\{a \sin 2\pi\left(\frac{m}{M} + \phi_f\right)\right\} \qquad m = 0, \ldots, M-1$$

Δ is the binwidth

w_m is a white-noise sequence with a Gaussian amplitude density, zero mean and unity variance.

$\mathrm{RND}[\cdot]$ is an operator which rounds its argument to the nearest integer.

With this implementation, the statistics of the period histogram given by equations A2.5a, A2.6, and A2.7 are correctly described

$$E[g_m] = s_m = R\lambda_o\Delta \exp\left\{a \sin 2\pi\left(\frac{m}{M} + \phi_f\right)\right\} \qquad \text{(A2.10a)}$$

$$\mathrm{var}\,[g_m] = s_m = E[g_m] \qquad \text{(A2.10b)}$$

The synchronization index $\hat{S}_f$ and $\hat{\phi}_f$ were the computed according to equations 3.52a and 3.52b.

For each value of a and N (the expected number of spikes in the period histogram), one hundred simulated histograms were formed according to equation A2.9. Each histogram consisted of 100 bins (i.e., M = 100). From the values of $(\hat{S}_f)_i$ and $(\hat{\phi}_f)_i$ computed from the i^{th} histogram, the mean values and variances of $\hat{S}_f$ and $\hat{\phi}_f$ were estimated according to

$$\hat{\bar{S}}_f = \frac{1}{100}\sum_{i=1}^{100}(\hat{S}_f)_i \qquad \text{(A2.11a)}$$

$$\hat{\sigma}^2_{\hat{S}_f} = \left[\frac{1}{100}\sum_{i=1}^{100}(\hat{S}_f)_i^2\right] - \hat{\bar{S}}_f^2 \qquad \text{(A2.11b)}$$

$$\hat{\bar{\phi}}_f = \frac{1}{100}\sum_{i=1}^{100}(\hat{\phi}_f)_i \qquad \text{(A2.11c)}$$

$$\hat{\sigma}^2_{\hat{\phi}_f} = \left[\frac{1}{100}\sum_{i=1}^{100}(\hat{\phi}_f)_i^2\right] - \hat{\bar{\phi}}_f^2 \qquad \text{(A2.11d)}$$

The results of this simulation are depicted in Figures A2.1 and A2.2. The estimate of the standard deviation of $\hat{S}_f$ is approximately inversely proportional to the square root of N. The general trend of this estimate is described by the empirical formula

$$\hat{\sigma}_{\hat{S}_f} = \frac{\left[\frac{\sqrt{4-\pi}}{2} + 0.25 \cdot \left[1 - \exp\{-10\, S_f\}\right]\right] \cdot \left[1 - S_f^2\right]}{\sqrt{N}} \qquad \text{(A2.12)}$$

The estimate of the standard deviation of the phase ϕ_f decreases with either larger N (i.e., more responses) or larger values of the synchronization index S_f. For $S_f = 0$, the standard deviation of ϕ_f is $\sqrt{3}/6 = 0.288$, equalling the standard deviation of a random variable uniformly distributed over (0, 1). No empirical formula for $\hat{\sigma}_{\hat{\phi}_f}$ was obtained.

Figure A2. 1. Estimate of the Standard Deviation of the Synchronization Index

Estimates of the standard deviation of $\hat{S}_f$, the measured value of S_f, are depicted as a function of S_f with N, the expected number of responses in the period histogram, as a parameter. Each data point is based upon the results of 100 period histogram simulations.

The vertical scale displays the standard deviation of $\hat{S}_f$ normalized to the square root of the number of responses. The empirical approximation depicted in the figure is given by:

$$\sqrt{N} \cdot \hat{\sigma}_{\hat{S}_f} = \left[\frac{\sqrt{4-\pi}}{2} + 0.25 \cdot (1 - \exp\{-10 \cdot S_f\})\right] \cdot \left[1 - S_f^2\right]$$

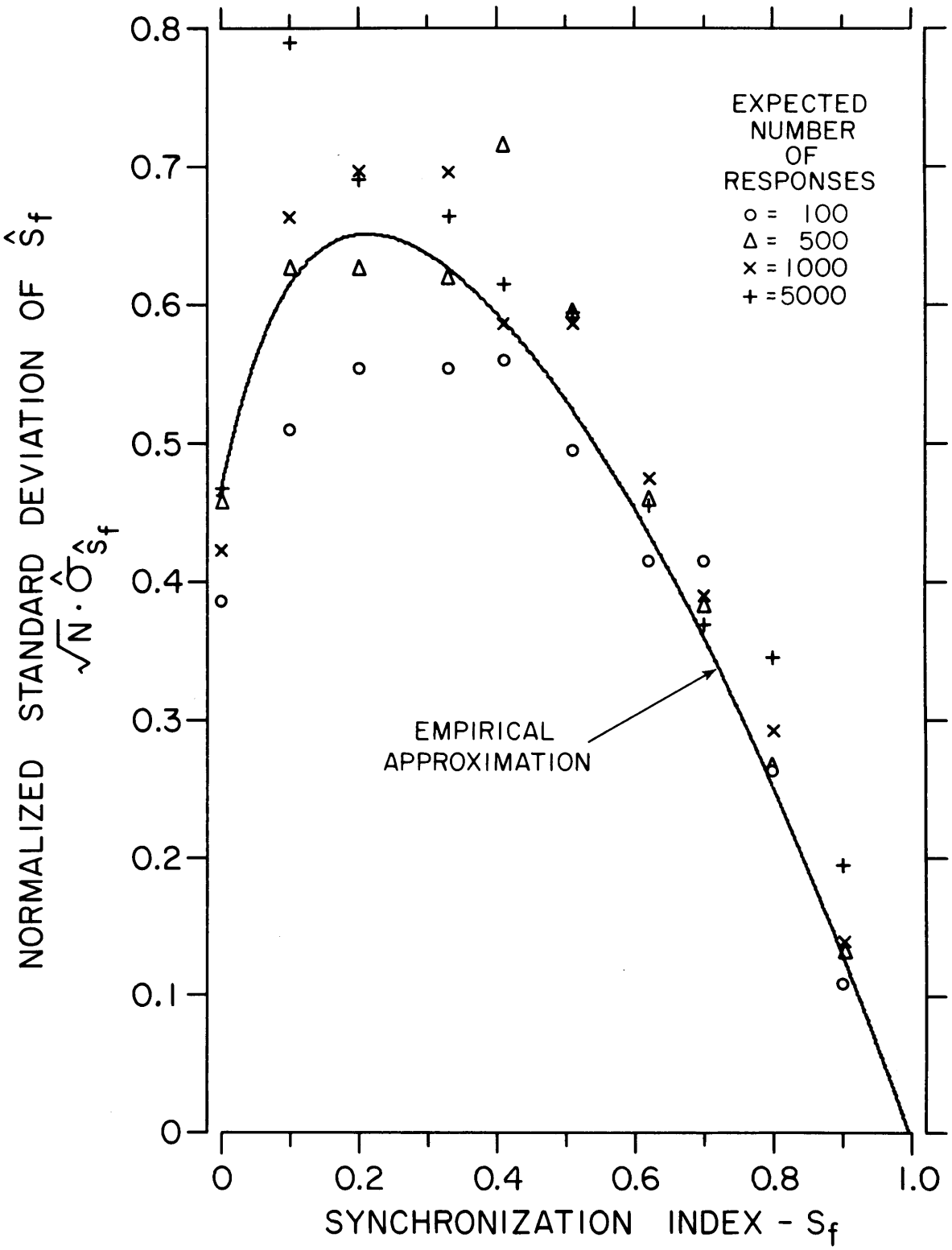
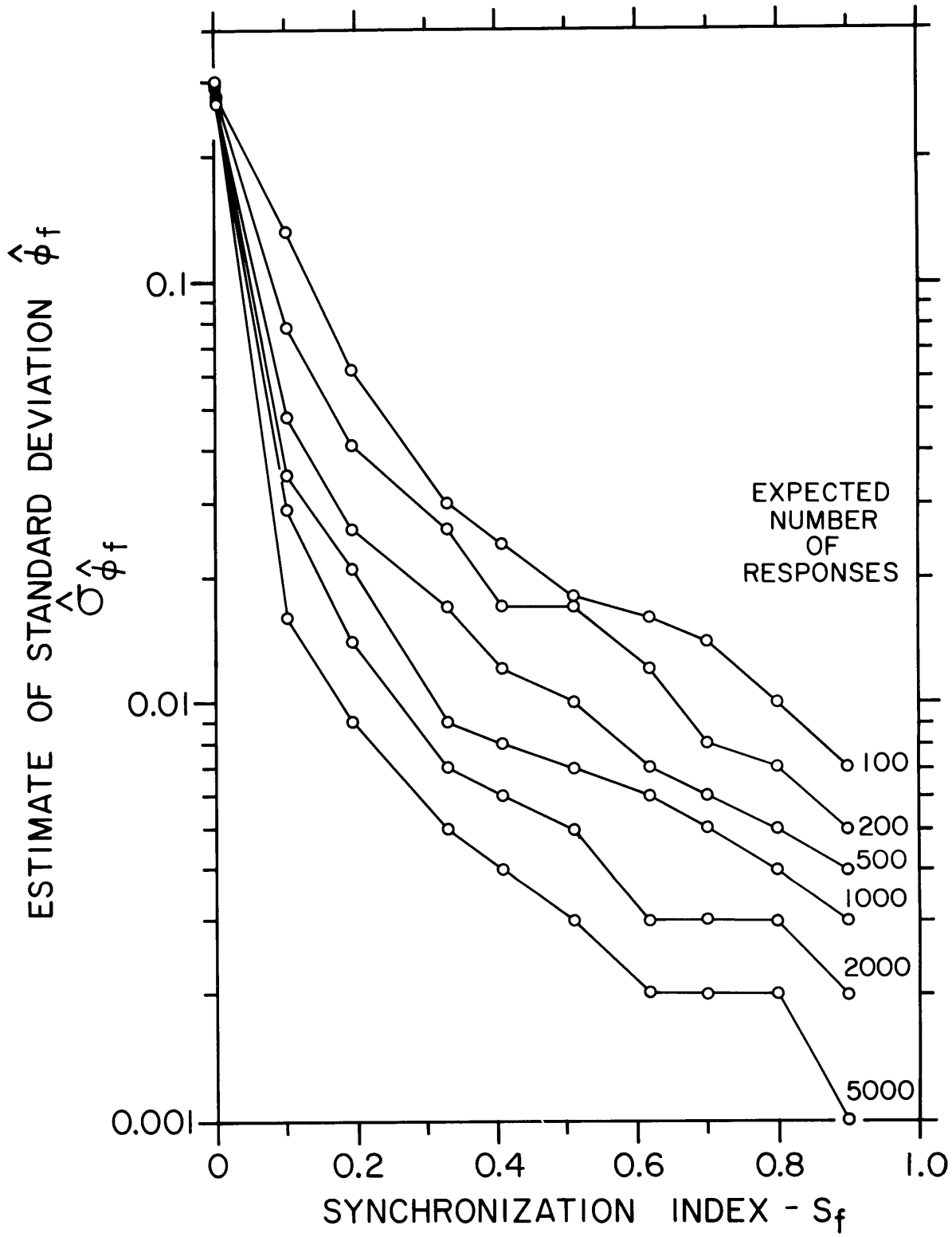

NORMALIZED STANDARD DEVIATION OF $\hat{S}_f$
$\sqrt{N} \cdot \hat{\sigma}_{\hat{S}_f}$
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.2
0.4
0.6
0.8
1.0
SYNCHRONIZATION INDEX - S_f
EXPECTED NUMBER OF RESPONSES
o = 100
Δ = 500
× = 1000
+ = 5000
EMPIRICAL APPROXIMATION

Figure A2. 2 Estimate of the Standard Deviation of the Phase of the Response

Estimates of the phase of the response are depicted as a function of S_f with the expected number of responses in the histogram as a parameter. Each data point is based upon the results of 100 period histogram simulations. The standard deviation of ϕ_f is plotted on a logarithmic scale.


ESTIMATE OF STANDARD DEVIATION $\hat{\phi}_f$
$\hat{\sigma}_{\hat{\phi}_f}$
0.1
0.01
0.001
0
0.2
0.4
0.6
0.8
1.0
SYNCHRONIZATION INDEX - S_f
EXPECTED NUMBER OF RESPONSES
100
200
500
1000
2000
5000

APPENDIX III

MAXIMUM LIKELIHOOD ESTIMATION OF PARAMETERS IN APPROXIMATIONS TO THE RATE OF DISCHARGE

This appendix is directed toward an exposition of the maximum likelihood estimation of the parameters of the expression given in equation 4.4

$$r(t) = \exp\{a_o + a_1 \cos 2\pi(ft+\theta_1) + a_2 \cos 2\pi(2ft-\theta_2)\} \qquad (4.4)$$

This estimation procedure has more generality than this specific problem. Some general results concerning the maximum likelihood estimation of the parameters of any approximations to a PST histogram are presented first. This application of maximum likelihood estimation was explicitly derived first by J. E. Evans in an unpublished work.

There are many techniques available for the estimation of signal parameters from a noisy version of the signal (e.g., Van Trees, 1968: 52-86). The maximum likelihood procedure provides estimates having the following properties (Cramér, 1946: 500-506):

1. The estimate is consistent; the estimate converges with probability one to the "correct value" as more data is obtained.
2. The estimate is asymptotically efficient: the variance of the estimate approaches the theoretical lower bound (the Cramér-Rao bound) as more data are included in the estimate.
3. The estimate is asymptotically normal (i.e., the pdf of the estimate approaches a Gaussian shape).

In order to derive the maximum likelihood estimation equations,

a probabilistic description of a PST histogram must be specified. A Bernoulli description is assumed here. Necessary conditions for the validity of this description are

1. No more than one spike may fall in any given bin during a particular stimulus presentation.
2. The probability of a spike falling into a bin is independent of the bins into which any previous spikes may have fallen.

These conditions are satisfied for period histograms (section 3.5).

Let g_m represent the contents of the m^{th} bin ($m = 0, \ldots, M-1$) of a PST histogram having a binwidth Δ. Define $P_m(\underline{a})$ to be the probability of a spike falling into bin m on any stimulus presentation. The probability $P_m(\underline{a})$ represents a characterization of the rate of discharge; the vector $\underline{a}$ is the L-component parameter vector of the characterization. Therefore

$$P_m(\underline{a}) = r_m(\underline{a}) \cdot \Delta \qquad \text{(A3.1)}$$

where

$r_m(\underline{a})$ is the rate of discharge as described by a given characterization.

Assuming the PST histogram is described by Bernoulli statictics; the probability of bin m containing the value g_m in R stimulus presentation is given by

$$\text{Prob}\,[g_m \mid \underline{a}] = \binom{R}{g_m} \cdot [P_m(\underline{a})]^{g_m} \cdot [1 - P_m(\underline{a})]^{R-g_m} \qquad \text{(A3.2)}$$

Define $P[\underline{g} \mid \underline{a}]$ to be the probability of obtaining a PST histogram having

the values $g_o, g_1, \ldots, g_{m-1}$ as given by the vector $\underline{g}$. The maximum likelihood estimate of the parameter vector $\underline{a}$ is defined to be the vector $\hat{\underline{a}}_{ML}$ which maximizes $P[\underline{g}|\underline{a}]$. Consequently, the vector $\hat{\underline{a}}_{ML}$ is computed by

$$\frac{\partial}{\partial a_i} P[\underline{g}|\underline{a}]\Big|_{\underline{a}=\hat{\underline{a}}_{ML}} = 0 \qquad i = 0, \ldots, L-1 \tag{A3.3a}$$

or, equivalently

$$\frac{\partial}{\partial a_i} \ln P[\underline{g}|\underline{a}]\Big|_{\underline{a}=\hat{\underline{a}}_{ML}} = 0 \qquad i = 0, \ldots, L-1 \tag{A3.3b}$$

Because of the assumption of Bernoulli statistics, the contents of distinct bins are independent random variables. The probability $P[\underline{g}|\underline{a}]$ is therefore written

$$P[\underline{g}|\underline{a}] = \prod_{m=0}^{M-1} \text{Prob}\,[g_m|\underline{a}] \tag{A3.4}$$

Substituting equation A3.4 into A3.3b, the result is

$$\sum_{m=0}^{M-1} \frac{\partial}{\partial a_i} \ln \left\{\text{Prob}\,[g_m|\underline{a}]\right\}\Big|_{\underline{a}=\hat{\underline{a}}_{ML}} = 0 \tag{A3.5}$$

Direct substitution of equation A3.2 into A3.5 does not yield a clean result. If, however, the conditions

$$R \gg g_m \tag{A3.6a}$$

$$P_m(\underline{a}) \ll 1 \tag{A3.6b}$$

are assumed to hold, then

$$R - g_m \approx R \tag{A3.7a}$$

$$\ln [1-P_m(\underline{a})] \approx -P_m(\underline{a}) \tag{A3.76b}$$

With these approximations, the substitution of equation A3.2 into A3.5 yields

$$\sum_{m=0}^{M-1} g_m \frac{\partial}{\partial a_i} [\ln P_m(\underline{a})] = R \sum_{m=0}^{M-1} \frac{\partial P_m(\underline{a})}{\partial a_i}\bigg|_{\underline{a}=\hat{\underline{a}}_{ML}} \qquad i=0,\ldots,L-1$$

(A3.8)

Expressing equation A3.8 in terms of $r_m(\underline{a})$, we have

$$\sum_{m=0}^{M-1} g_m \frac{\partial \ln r_m(\underline{a})}{\partial a_i} = R\Delta \sum_{m=0}^{M-1} \frac{\partial r_m(\underline{a})}{\partial a_i}\bigg|_{\underline{a}=\hat{\underline{a}}_{ML}} \qquad i = 0,\ldots,L-1$$

(A3.9)

In general, for an arbitrary $r_m(\underline{a})$, the solution to this set of L simultaneous equations is the maximum likelihood estimate of the parameter vector $\underline{a}$. Next we shall consider the special case for which the rate of discharge is given by equation 4.4. Rewriting this equation in terms of $r_m(\underline{a})$, we have

$$r_m(\underline{a}) = \exp\left\{a_o + a_1 \cos 2\pi\left(\frac{m}{M} + \theta_1\right) + a_2 \cos 2\pi\left(\frac{2m}{M} + \theta_2\right)\right\} \qquad m = 0,\ldots,M-1$$

(A3.10)

where

f, the fundamental frequency, is given by $f = 1/M\Delta$.

The maximum likelihood equations for this characterization are

for $a_i = a_o$:
$$\sum_{m=0}^{M-1} g_m = R\Delta \sum_{m=0}^{M-1} r_m(\underline{a}) \tag{A3.11a}$$

for $a_i = a_1$:
$$\sum_{m=0}^{M-1} g_m \cos 2\pi\left(\frac{m}{M} + \theta_1\right) = R\Delta \sum_{m=0}^{M-1} \cos 2\pi\left(\frac{m}{M} + \theta_1\right) r_m(\underline{a}) \tag{A3.11b}$$

for $a_i = \theta_1$:
$$\sum_{m=0}^{M-1} g_m \sin 2\pi\left(\frac{m}{M} + \theta_1\right) = R\Delta \sum_{m=0}^{M-1} \sin 2\pi\left(\frac{m}{M} + \theta_1\right) \cdot r_m(\underline{a}) \tag{A3.11c}$$

for $a_i = a_2$:
$$\sum_{m=0}^{M-1} g_m \cos 2\pi\left(\frac{2m}{M} + \theta_2\right) = R\Delta \sum_{m=0}^{M-1} \cos 2\pi\left(\frac{2m}{M} + \theta_2\right) \cdot r_m(\underline{a}) \tag{A3.11d}$$

for $a_i = \theta_2$:
$$\sum_{m=0}^{M-1} g_m \sin 2\pi\left(\frac{2m}{M} + \theta_2\right) = R\Delta \sum_{m=0}^{M-1} \sin 2\pi\left(\frac{2m}{M} + \theta_2\right) \cdot r_m(\underline{a}) \tag{A3.11e}$$

Define the quadrature components of S_f and S_{2f}

$$S_{f,c} = \frac{\sum_{m=0}^{M-1} g_m \cos \frac{2\pi m}{M}}{\sum_{m=0}^{M-1} g_m} \tag{A3.12a}$$

$$S_{f,s} = \frac{\sum_{m=0}^{M-1} g_m \sin \frac{2\pi m}{M}}{\sum_{m=0}^{M-1} g_m} \tag{A3.12b}$$

$$S_{2f,\,c} = \frac{\sum_{m=0}^{M-1} g_m \cos \frac{4\pi m}{M}}{\sum_{m=0}^{M-1} g_m} \qquad \text{(A3.12c)}$$

$$S_{2f,\,s} = \frac{\sum_{m=0}^{M-1} g_m \sin \frac{4\pi m}{M}}{\sum_{m=0}^{M-1} g_m} \qquad \text{(A3.12d)}$$

The equations A3.11a-e become

$$\sum_{m=0}^{M-1} g_m = R\Delta \sum_{m=0}^{M-1} r_m(\underline{a}) \qquad \text{(A3.13a)}$$

$$\sum_{m=0}^{M-1} g_m \cdot \left(S_{f,\,c} \cos 2\pi\theta_1 - S_{f,\,s} \sin 2\pi\theta_1\right) = R\Delta \sum_{m=0}^{M-1} \cos 2\pi\left(\frac{m}{M}+\theta_1\right) r_m(\underline{a}) \qquad \text{(A3.13b)}$$

$$\sum_{m=0}^{M-1} g_m \cdot \left(S_{f,\,c} \sin 2\pi\theta_1 + S_{f,\,s} \cos 2\pi\theta_1\right) = R\Delta \sum_{m=0}^{M-1} \sin 2\pi\left(\frac{m}{M}+\theta_1\right) r_m(\underline{a}) \qquad \text{(A3.13c)}$$

$$\sum_{m=0}^{M-1} g_m \cdot \left(S_{2f,\,c} \cos 2\pi\theta_2 - S_{2f,\,s} \sin 2\pi\theta_2\right) = R\Delta \sum_{m=0}^{M-1} \cos 2\pi\left(\frac{2m}{M}+\theta_2\right) r_m(\underline{a}) \qquad \text{(A3.13d)}$$

$$\sum_{m=0}^{M-1} g_m \cdot \left(S_{2f,\,c} \sin 2\pi\theta_2 + S_{2f,\,s} \cos 2\pi\theta_2\right) = R\Delta \sum_{m=0}^{M-1} \sin 2\pi\left(\frac{2m}{M}+\theta_2\right) r_m(\underline{a}) \qquad \text{(A3.13e)}$$

Hence the maximum likelihood estimates of the parameters of the exponential approximation are determined from the period histogram by specifying the quantities $S_{f,c}$, $S_{f,s}$, $S_{2f,c}$, $S_{2f,s}$, and $\sum_{m=0}^{M-1} g_m$. These are termed the sufficient statistics of the estimation procedure: these are the only measures of the data that are required to obtain estimates of the parameters in equation 4.4. In addition to $\bar{r}$, S_f, and ϕ_f, the parameters needed are the synchronization index S_{2f} and the phase ϕ_{2f} evaluated at the second harmonic of the fundamental. The sufficient statistics listed above are derived from these measurements.

Using the fact that (Abramowitz and Stegun 1965: 376)

$$\exp\{a\cos\alpha\} = I_o(a) + 2\sum_{k=1}^{\infty} I_k(a)\cos k\alpha, \qquad \text{(A3.14)}$$

where

$I_k(a)$ denotes the k^{th} order modified Bessel function of the first kind ($k = 0, \ldots$).

Equation A3.13a may be expressed

$$\sum_{m=0}^{M-1} g_m = MR\Delta \cdot \exp\{a_o\} \cdot A \qquad \text{(A3.15)}$$

where

A is defined to be the expression

$$A \equiv I_o(a_1)\, I_o(a_2) + 2 \sum_{k=1}^{\infty} I_{2k}(a_1)\, I_k(a_2) \cos 2\pi k(2\theta_1 - \theta_2) \tag{A3.16}$$

Therefore, the solution for the parameter a_o is given by

$$a_o = \ln\left[\frac{\sum_{m=0}^{M-1} h_m}{MR\Delta \cdot A}\right] \tag{A3.17}$$

To evaluate A requires the solution of equation A3.17 for the parameters a_1, θ_1, a_2, and θ_2. Explicitly writing these equations, making extensive use of equations A3.10 and A3.14

$$A \cdot \left[S_{f,c} \cos 2\pi\theta_1 - S_{f,s} \sin 2\pi\theta_1\right] =$$

$$I_1(a_1) I_o(a_2) + \sum_{k=1}^{\infty} I_k(a_2) \cdot \left[I_{2k-1}(a_1) + I_{2k+1}(a_1)\right] \cos 2\pi k(2\theta_1 - \theta_2) \tag{A3.18a}$$

$$A \cdot \left[S_{f,c} \sin 2\pi\theta_1 + S_{f,s} \cos 2\pi\theta_1\right] = \sum_{k=1}^{\infty} \left[I_k(a_2) \cdot \left[I_{2k+1}(a_1) - I_{2k-1}(a_1)\right] \cdot \sin 2\pi k(2\theta_1 - \theta_2)\right] \tag{A3.18b}$$

$$A \cdot \left[S_{2f,\,c} \cos 2\pi\theta_2 - S_{2f,\,s} \sin 2\pi\theta_2\right] =$$

$$I_2(a_1)\, I_o(a_2) \cos 2\pi(2\theta_1 - \theta_2)$$

$$+ \sum_{k=1}^{\infty} I_k(a_2) \cdot \left[I_{2k-2}(a_1) \cos\left\{2\pi(k-1)(2\theta_1 - \theta_2)\right\}\right.$$

$$\left. + I_{2k+2}(a_1) \cos\left\{2\pi(k+1)(2\theta_1 - \theta_2)\right\}\right] \qquad \text{(A3.18c)}$$

$$A \cdot \left[S_{2f,\,c} \sin 2\pi\theta_2 + S_{2f,\,s} \cos 2\pi\theta_2\right] =$$

$$I_2(a_1)\, I_o(a_2) \sin 2\pi(2\theta_1 - \theta_2)$$

$$+ \sum_{k=1}^{\infty} I_k(a_2) \cdot \left[I_{2k+2}(a_1) \sin\left\{2\pi(k+1)(2\theta_1 - \theta_2)\right\}\right.$$

$$\left. - I_{2k-2}(a_1) \cdot \sin\left\{2\pi(k-1)(2\theta_1 - \theta_2)\right\}\right] \qquad \text{(A3.18d)}$$

These equations must be solved for the parameters a_1, θ_1, a_2, and θ_2.

The solution of these four, simultaneous, nonlinear equations was accomplished by a four-dimensional version of the Newton-Raphson root-calculation procedure (Hildebrand, 1956: 451). This iterative procedure reduces the nonlinear equations to a set of linear

equations, the solutions to which are corrections to trial values of the parameters. The new parameters are substituted and new corrections are computed; this procedure is iterated until the corrections are nearly zero.

The program implementing this procedure was written in the interpretive language FOCAL on a Digital Equipment Corporation PDP-8/E computer. FOCAL performs computations in floating point with a mantissa length of 24 bits. The iterative procedure outlined above did not converge in 4 out of 1050 histograms. The procedure can be sensitive to the initial guess of the values of the parameters. In the non-convergent cases, several initial values were used, but none of these resulted in a convergent solution.

BIOGRAPHICAL NOTE

Born:

July 9, 1946
Mt. Pleasant, Texas

Education:

S. B. and S. M.	Electrical Engineering, Massachusetts Institute of Technology - 1970
E. E.,	Electrical Engineering, Massachusetts Institute of Technology - 1971

Research and/or Professional Experience:

1971 - present	Instructor in Electrical Engineering Massachusetts Institute of Technology
1971 - present	Research Assistant in Otolaryngology, Eaton-Peabody Laboratory Massachusetts Eye and Ear Infirmary
1970-1971	Teaching Assistant Department of Electrical Engineering Massachusetts Institute of Technology
1968-1973	Summer Staff Member Lincoln Laboratory Lexington, Massachusetts

Publications:

Oppenheim, A., and Johnson, D. (1972). "Discrete representation of signals," Proc. IEEE, 60 (6): 681-691

Oppenheim, A., Johnson, D., and Steiglitz, K. (1971). "Computation of Spectra with Unequal Resolution Using the Fast Fourier Transform," Proc. IEEE (Letter), 59 (2): 299-301.

Oppenheim, A. V., and Johnson, D. H. (1970). "Discrete Representation of Analog Signals," M. I. T. Research Laboratory of Electronics, Quarterly Progress Report No. 97: 185-190.

Johnson, D. H. (1970). "Statistical Relationships Between the Firing Patterns of Two Auditory Nerve Fibers," M. I. T. Research Laboratory of Electronics, Quarterly Progress Report No. 97: 173-180.